ON THE LEFSCHETZ FIXED POINT FORMULA.

By Robert F. Brown.[1]

Let X be the closed triangulable n-manifold, i.e., a compact, connected, metric, n-dimensional, locally Euclidean space which is homeomorphic to a finite polyhedron, and let $f\colon X\to X$ be a map with a finite number of fixed points $x_1,\cdots,x_r$. Lefschetz [6] defined the index $i_f(x_j)\in Z$ (the integers) of the fixed points x_j, $j=1,\cdots,r$, and proved the "Lefschetz formula"

$$\sum_{j=1}^{r} i_f(x_j) = (-1)^n \Lambda_f$$

where Λ_f is the Lefschetz number of f.[2] The Lefschetz formula was subsequently verified by Lefschetz [7] for compact triangulated manifolds with boundary, generalized by Leray [8] to convexoid spaces, and generalized by Browder [1] to semi-complexes. All these results use definitions of index based on induced chain mappings. A definition of index for isolated fixed points which depends on the notion of the degree of a map on a sphere was given by Hopf [5] and the Lefschetz formula for maps on finite polyhedra was proved using this definition.

An immediate corollary to the Lefschetz formula is the following: if X is a closed triangulable n-manifold, $f,f'\colon X\to X$ are maps with fixed points $x_1,\cdots,x_r$ and $x'_1,\cdots,x'_s$, respectively, and f is homotopic to f', then

$$\sum_{j=1}^{r} i_f(x_j) = \sum_{k=1}^{s} i_{f'}(x'_k).$$

Weier [9] proved this corollary for maps on topological (i.e., not necessarily triangulable) manifolds using the degree index. However, for this definition of index, no proof of the Lefschetz formula itself exists for closed topological manifolds. We shall define a fixed point index $I_f(x)$ which is equivalent to the one defined by Weier and prove the following result:

Received November 22, 1963.
Revised September 18, 1964.

[1] This research was supported in part by the Air Force Office of Scientific Research under grant AFOSR 90-63.

[2] The term $(-1)^n$ does not appear explicitly in the Lefschetz formula of [6], but is included in the definition of the index. We are here using Hopf's formulation of the result [5].

THEOREM. *Let M be a closed orientable (in the sense of Fadell [4]) topological n-manifold and let $f\colon M \to M$ be a map with fixed points $x_1, \cdots, x_r$, then*

$$\sum_{j=1}^{r} I_f(x_j) = (-1)^n \Lambda_f.$$

We shall use singular cohomology theory with integer coefficients. If X and Y are topological spaces with $A \subseteq X$, $B \subseteq Y$, then $f\colon (X,A) \to (Y,B)$ is a *map* if f is a map of X into Y and $f(A) \subseteq B$. Those homomorphisms of cohomology groups which are not otherwise identified will be assumed to be induced by inclusions. The term *n-manifold* will here mean closed orientable topological n-manifold.

Let M be an n-manifold and let x_j be an isolated fixed point of a map $f\colon M \to M$, i.e., there is a Euclidean neighborhood U_j of x_j containing no other fixed point of f. Let

$$\Delta_j = \{(y,z) \in U_j \times M \mid y = z\}$$

and define

$$k_j\colon (M, M - x_j) \to (U_j \times M, U_j \times M - \Delta_j)$$

by $k_j(y) = (x_j, y)$. Let $p\colon U_j \times M \to U_j$ be projection, then

$$(U_j \times M, U_j \times M - \Delta_j, p, U_j)$$

is a fibred pair [4, 3.5] and since U_j is orientable, it follows from the proof of Theorem 5.2 of [4] that

$$k_j^*\colon H^n(U_j \times M, U_j \times M - \Delta_j) \to H^n(M, M - x_j)$$

is an isomorphism. Define

$$(1 \times f)_j\colon (U_j, U_j - x_j) \to (U_j \times M, U_j \times M - \Delta_j)$$

by $(1 \times f)_j(y) = (y, f(y))$. Let the homomorphism $F^*{}_j$ be defined so that the diagram

$$\begin{array}{ccc}
H^n(M, M - x_j) & \xrightarrow{\quad F^*{}_j \quad} & H^n(M, M - x_j) \\
k_j^* \Big\uparrow \approx & & \Big\downarrow \approx \\
H^n(U_j \times M, U_j \times M - \Delta_j) & \xrightarrow[(1 \times f)_j^*]{} & H^n(U_j, U_j - x_j)
\end{array}$$

commutes. For a generator $\mu^*{}_j \in H^n(M, M - x_j)$ we define the *index* of f

at x_j, $I_f(x_j)$, by $F^*_j(\mu^*_j) = I_f(x_j)\cdot\mu^*_j$. Since a discussion of the necessary technical questions which arise from the definition of a fixed point index would interrupt the proof of the main result, we will postpone this discussion until the end of the paper.

When $f\colon M\to M$ is a map with fixed points $x_1,\cdots,x_r$, we let $\{U_j\}$, $j=1,\cdots,r$, be disjoint Euclidean neighborhoods of the x_j such that for $y\in U_j$, $f(y)=y$ if and only if $y=x_j$. Pick a generator $\mu\in H^n(M)\cong Z$ and let μ^*_j be the pre-image of μ with respect to the isomorphism

$$H^n(M,M-x_j)\to H^n(M).$$

For $a\in Z$, define

$$\beta\colon H^n(M)\to\sum_{j=1}^{r}H^n(M,M-x_j)$$

by $\beta(a\cdot\mu)=(a\cdot\mu^*_1,\cdots,a\cdot\mu^*_r)$. From the exact cohomology sequence of the pair (U_j,U_j-x_j), we obtain the isomorphism

$$\delta^*_j\colon H^{n-1}(U_j-x_j)\to H^n(U_j,U_j-x_j).$$

Let $\mu'_j\in H^n(U_j,U_j-x_j)$ be the generator which is the pre-image of μ^*_j with respect to the isomorphism

$$H^n(M,M-x_j)\to H^n(U_j,U_j-x_j)$$

and set $\mu''_j=\delta^*_j{}^{-1}(\mu'_j)$. Let $h\colon S^{n-1}\to U_1-x_1$ be an embedding and set $\nu=h^*(\mu''_1)\in H^{n-1}(S^{n-1})$. A typical element of $\sum_{j=1}^{r}H^{n-1}(U_j-x_j)$ is of the form $(a_1\cdot\mu''_1,\cdots,a_r\cdot\mu''_r)$ where $a_1,\cdots,a_r\in Z$. We define

$$\gamma\colon\sum_{j=1}^{r}H^{n-1}(U_j-x_j)\to \mathrm{H}^{n-1}(S^{n-1})$$

by $\gamma(a_1\cdot\mu''_1,\cdots,a_r\cdot\mu''_r)=(a_1+\cdots+a_r)\cdot\nu$. Now consider the endomorphism A^* of $H^*(M)$ defined so that diagram (1) commutes where k^*, $(1\times f)^*$, and δ^* are obtained from the k_j^*, $(1\times f)_j^*$, and δ^*_j respectively in the obvious way. The proof that $A^*(\mu)=\sum_{j=1}^{r}I_f(x_j)\cdot\mu$ is entirely straightforward and we therefore omit it.

The proof of a theorem of M. Brown and B. Casler [2] can be modified slightly to obtain the following result. Given an n-manifold M and points $x_1,\cdots,x_r$ in M there exists a map of the n-cell C_n onto M such that the restriction of the map to the interior of C_n is a homeomorphism whose image contains $x_1,\cdots,x_r$. Hence we can assume that the fixed points of f are

in a Euclidean neighborhood (homeomorphic image of Euclidean n-dimensional space). Call the neighborhood U. We may also assume that the disjoint Euclidean neighborhoods U_j of the x_j are contained in U.

$$\begin{tikzcd}
H^n(M) \arrow[r, "A^*"] \arrow[d, "\beta"'] & H^n(M) \\
\sum_{j=1}^{r} H^n(M, M-x_j) & H^n(M, M-x_1) \arrow[u, "\approx"] \arrow[d, "\approx"'] \\
\sum_{j=1}^{r} H^n(U_j \times M, U_j \times M - \Delta_j) \arrow[u, "k^*"', "\approx"] \arrow[d, "(1\times f)^*"'] & H^n(U_1, U_1 - x_1) \\
\sum_{j=1}^{r} H^n(U_j, U_j - x_j) & H^{n-1}(U_1 - x_1) \arrow[u, "\approx", "\delta^*_1"'] \arrow[d, "\approx"', "h^*"] \\
\sum_{j=1}^{r} H^{n-1}(U_j - x_j) \arrow[u, "\delta^*"'] \arrow[r, "\gamma"'] & H^{n-1}(S^{n-1})
\end{tikzcd}$$

DIAGRAM (1)

Define $d\colon M \to M \times M$ to be the diagonal map, let $\Delta = d(M)$, and define $\bar{f}\colon M \times M \to M \times M$ by $\bar{f}(y,z) = (y, f(z))$. Let λ be the endomorphism of $H^n(M)$ defined so that Diagram (2) commutes, then by Theorem 2.2 of [3], $\lambda(\mu) = (-1)^n \Lambda_f \cdot \mu$. Note that in Diagram (2), the homomorphism $k_1{}^*$ is an isomorphism by the argument used for $k_j{}^*$ above because M is orientable.

$$\begin{tikzcd}
H^n(M) \arrow[r, "\lambda"] & H^n(M) \\
H^n(M, M - x_1) \arrow[u, "\approx"'] & H^n(M \times M) \arrow[u, "d^*"'] \\
H^n(M \times M, M \times M - \Delta) \arrow[u, "k_1{}^*", "\approx"'] \arrow[r] & H^n(M \times M) \arrow[u, "\bar{f}^*"']
\end{tikzcd}$$

DIAGRAM (2)

Define $\lambda'\colon H^n(M) \to H^n(M)$ so that Diagram (3) commutes where $\Delta_0 = \Delta \cap (U \times M)$, $\phi_r = \{x_1, \cdots, x_r\}$, and $(1 \times f)(y) = (y, f(y))$.

$$\begin{tikzcd}
H^n(M) \arrow[rr, "\lambda'"] & & H^n(M) \\
H^n(M, M - x_1) \arrow[u, "\approx"'] & & H^n(M, M - \phi_r) \arrow[u] \arrow[d, "\approx"'] \\
H^n(U \times M, U \times M - \Delta_0) \arrow[u, "k_1^*", "\approx"'] \arrow[rr, "(1 \times f)^*"'] & & H^n(U, U - \phi_r)
\end{tikzcd}$$

DIAGRAM (3)

$$\begin{tikzcd}
 & H^n(M) & \\
 & H^n(M, M - x_1) \arrow[u] & \\
H^n(M \times M) \arrow[d, "\bar{f}^*"'] & H^n(M \times M, M \times M - \Delta) \arrow[l] \arrow[r] \arrow[u, "k_1^*", "\approx"'] \arrow[dd, "(1 \times f)^*"] & H^n(U \times M, U \times M - \Delta_0) \arrow[ul, "\approx", "k_1^*"'] \arrow[d, "(1 \times f)^*"] \\
H^n(M \times M) \arrow[d, "d^*"'] & & H^n(U, U - \phi_r) \\
H^n(M) \arrow[drr] & H^n(M, M - \phi_r) \arrow[l] \arrow[r] & H^n(M, M - \phi_r) \arrow[u, "\approx"'] \arrow[d] \\
 & & H^n(M)
\end{tikzcd}$$

DIAGRAM (4)

The fact that $\lambda'(\mu) = (-1)^n \Delta_f \cdot \mu$ follows from the commutativity of Diagram (4) since λ is the composition of the homomorphisms—or their inverses—on the left side of the diagram and λ' is obtained in the same way from the right side.

$$
\begin{tikzcd}
 & & H^n(M) \arrow[dll, "\beta"'] \arrow[drr, "\approx"] & & \\
\sum_{j=1}^{r} H^n(M, M-x_j) \arrow[rr] & & \sum_{j=1}^{r} H^n(U_j, U_j-x_j) \arrow[r] & H^n(U_1, U_1-x_1) & H^n(M, M-x_1) \arrow[l] \\
 & & H^n(U_1\times M, U_1\times M-\Delta_1) \arrow[urr] & & \\
\sum_{j=1}^{r} H^n(U_j\times M, U_j\times M-\Delta_j) \arrow[uu, "k^*", "\approx"'] \arrow[urr] \arrow[d, "(1\times f)^*"'] & & & & H^n(U\times M, U\times M-\Delta) \arrow[llll] \arrow[uu, "\approx", "k_1^*"'] \arrow[d, "(1\times f)^*"] \\
\sum_{j=1}^{r} H^n(U_j, U_j-x_j) \arrow[dd, "\gamma'"'] & & & & H^n(U, U-\phi_r) \arrow[llll, "\approx"'] \arrow[dll, "j^*"'] \arrow[ddd, "\approx"', "l^*"] \\
 & (6) & \sum_{j=1}^{r} H^n(U_j, U_j-x_j) \arrow[dll, "\gamma''"] & & \\
H^n(U_1, U_1-x_1) & & & (7) & \\
H^n(M, M-x_1) \arrow[u, "i_1^*", "\approx"'] \arrow[drr, "k_1^*"'] & & & & H^n(M, M-\phi_r) \arrow[dll, "\bar{k}^*"] \\
 & & H^n(M) & &
\end{tikzcd}
$$

DIAGRAM (5)

Consider Diagram (5) where $\gamma' = \delta_1^{*-1} h^{*-1} \gamma \delta^*$ (see Diagram (1)) and γ'' is defined to make subdiagram (6) commute. The homomorphism A^* is obtained from the left side of Diagram (5) and λ' is obtained from the right side. *Therefore, in order to prove the Lefschetz formula, it is sufficient to prove that Diagram* (5) *commutes.*

Diagram (5) will commute if every subdiagram of it does. The only difficult step is the proof of the commutativity of subdiagram (7) (where i^*, j^*, $\bar{k}^*$, $i_1{}^*$, and $\bar{k}_1{}^*$ are inclusion induced homomorphisms). The commutativity of subdiagram (7) has nothing to do with the manner in which the set ϕ_r was obtained. Therefore, we may use induction on r, the number of points. In the case $r=1$, we have Diagram (8) and since $ij=i_1$, we are through.

$$\begin{tikzcd}
H^n(U_1, U_1 - x_1) & & H^n(U_1, U_1 - x_1) \arrow[ll, "\gamma'' = \mathrm{id.}"'] \\
H^n(M, M - x_1) \arrow[u, "i_1{}^*"] \arrow[d, "\bar{k}_1{}^*"', "\approx"] \arrow[rr, "\approx", "i^*"'] & & H^n(U, U - x_1) \arrow[u, "j^*"'] \\
H^n(M) & & H^n(M, M - x_1) \arrow[u, "i^*"'] \arrow[ll, "\bar{k}^* = \bar{k}_1{}^*"]
\end{tikzcd}$$

DIAGRAM (8)

Let $\gamma^*{}_r \colon H^n(M, M - \phi_r) \to H^n(M)$ be defined by $\gamma^*{}_r = \bar{k}_1{}^* i_1{}^{*-1} \gamma'' j^* i^*$ and let

$$q \colon (M, M - \phi_r) \to (M, M - \phi_{r-1})$$

be the inclusion where $\phi_{r-1} = \{x_1, \cdots, x_{r-1}\}$. In Diagram (9), the commutativity of subdiagram (10) follows from the induction hypothesis, so it remains to show that subdiagram (11) commutes.

Let $(\mu'_1, \cdots, \mu'_r)$ generate $\sum_{j=1}^{r} H^n(U_j, U_j - x_j)$, then

$$(\nu_1, \cdots, \nu_r) = i^{*-1} j^{*-1} (\mu'_1, \cdots, \mu'_r)$$

generates $H^n(M, M - \phi_r)$. For $a_1, \cdots, a_{r-1} \in Z$ we have

$$q^* (a_1 \cdot \nu_1, \cdots, a_{r-1} \cdot \nu_{r-1}) = (a_1 \cdot \nu_1, \cdots, a_{r-1} \cdot \nu_{r-1}, 0 \cdot \nu_r).$$

Also, for $a'_1, \cdots, a'_r \in Z$

$$\gamma^*{}_r (a'_1 \cdot \nu_1, \cdots, a'_r \cdot \nu_r) = (a'_1 + \cdots + a'_r) \cdot \mu.$$

Thus $\gamma^*{}_r q^* = \gamma^*{}_{r-1}$ and Diagram (11) commutes which shows that the induction works. Therefore Diagram (7) commutes, Diagram (5) commutes, and the theorem is proved.

There remain two technical points which must be established in connection with the theorem we have just proved. We must, of course, show that

the index $I_f(x)$ is uniquely defined. Also, in order to prove that our result is both an improvement on Weier's and a true "Lefschetz formula," we will show that the index $I_f(x)$ is equivalent to the index used by Weier and hence it is a true generalization of the index defined by Lefschetz.

$$
\begin{tikzcd}
H^n(M, M-\phi_r) \arrow[rr, "\gamma^*_r"] & & H^n(M) \\
& H^n(M, M-\phi_{r-1}) \arrow[ul, "q^*"'] \arrow[r, "\gamma^*_{r-1}"] \arrow[ur, phantom, "(11)"] & H^n(M) \arrow[u] \\
& H^n(M, M-\phi_{r-1}) \arrow[u] \arrow[r, "k^*"'] \arrow[dl, "q^*"'] \arrow[ur, phantom, "(10)"] & H^n(M) \arrow[u] \\
H^n(M, M-\phi_r) \arrow[uuu] \arrow[rr, "k^*"'] & & H^n(M) \arrow[u]
\end{tikzcd}
$$

DIAGRAM (9)

In order to show that $I_f(x_j)$ is well-defined, we observe that it is clearly independent of the choice of the generator of $H^n(M, M-x_j)$ so that we need only show that it is independent of the choice of the Euclidean neighborhood U_j. Let V_j be any other Euclidean neighborhood of x_j and let $W \subset U_j \cap V_j$ be a Euclidean neighborhood of x_j. The obvious commutativity of the diagram

$$
\begin{tikzcd}
H^n(U_j \times M, U_j \times M - \Delta_j) \arrow[r, "(1\times f)_j^*"'] \arrow[d] & H^n(U_j, U_j - x_j) \arrow[d, "\approx"'] \\
H^n(W \times M, W \times M - \Delta_W) \arrow[r, "(1\times f)_W^*"'] & H^n(W, W - x_j)
\end{tikzcd}
$$

where $(1 \times f)_W = (1 \times f)_j \mid W$ and $\Delta_W = \Delta_j \cap (W \times M)$, proves that the index defined using W in place of U_j is the same as $I_f(x_j)$. Since the identical argument shows that the index defined using V_j in place of U_j is the same as that using W, we are through.

The index used by Weier, which we will write as $i_f(x)$, may be defined

as follows. Let E^n denote Euclidean n-dimensional space and let $h\colon E^n \to U_j$ be a homeomorphism (onto) taking the origin to x_j. Since x_j is a fixed point of f, there is an n-cell

$$B = \{z \in E^n \mid |z| \leqq r \neq 0\}$$

(where $|z|$ is the distance from z to the origin in the Euclidean metric) such that $fh(B) \subset U_j$. Define $g\colon B \to E^n$ by $g(z) = h^{-1}fh(z) - z$. Let ρ be the retraction of E^n to B obtained by setting $\rho(z) = r \cdot |z|^{-1} \cdot z$ for $z \in E^n - B$. Now define $G\colon (B, B - 0) \to (B, B - 0)$ by $G = \rho g$, then G induces an endomorphism G^* of $H^n(B, B - 0)$. Let $\eta \in H^n(B, B - 0) \cong Z$ be a generator and define the index $i_f(x_j)$ by $G^*(\eta) = i_f(x_j) \cdot \eta$.

Fadell [4] proved that the map

$$\zeta\colon (U_j \times M, U_j \times M - \Delta_j) \to (U_j \times M, U_j \times (M - x_j))$$

given by

$$\zeta(y, z) = \begin{cases} (y, z) & z \in M - U_j \\ (y, h(h^{-1}(z) - h^{-1}(y))) & z \in U_j \end{cases}$$

is a homeomorphism.

$$\begin{tikzcd}
H^n(M, M - x_j) & & H^n(M, M - x_j) \arrow[ll] \\
H^n(U_j \times M, U_j \times M - \Delta_j) \arrow[u, "k_j^*"', "\approx"] \arrow[ddd, "(1 \times f)^*"'] & & H^n(U_j \times M, U_j \times (M - x_j)) \arrow[ll, "\approx", "\zeta^*"'] \arrow[u, "\approx", "k_j^*"'] \arrow[d, "\zeta^*"] \\
& & H^n(U_j \times M, U_j \times M - \Delta_j) \arrow[dd, "(1 \times f)^*"] \\
& & \\
H^n(U_j, U_j - x_j) & & H^n(U_j, U_j - x_j) \arrow[ll] \\
H^n(M, M - x_j) \arrow[u, "\approx"'] & & H^n(M, M - x_j) \arrow[ll] \arrow[u, "\approx", "i^*"']
\end{tikzcd}$$

DIAGRAM (12)

If we observe that $\zeta(x_j, z) = (x_j, z)$ for all $z \in M$, then it is obvious that Diagram (12) commutes and we have

$$I_f(x_j) \cdot \mu^*_j = i^{*-1}(1 \times f)^* \zeta^* k_j^{*-1}(\mu^*_j).$$

Consider Diagram (13) where $s\colon B \to E^n$ is inclusion.

$$\begin{tikzcd}
(M, M - x_j) \arrow[d, "k_j"'] & (U_j, U_j - x_j) \arrow[l] & (B, B - 0) \arrow[l, "h \mid B"'] \arrow[d, "s"] \\
(U_j \times M, U_j \times (M - x_j)) & (U_j, U_j - x_j) \arrow[l, "k_j \mid U_j"] & (E^n, E^n - 0) \arrow[l, "h"] \\
(U_j \times M, U_j \times M - \Delta_j) \arrow[u, "\zeta"] & & \\
(U_j, U_j - x_j) \arrow[u, "(1 \times f)"] \arrow[d, "i"'] & & (B, B - 0) \arrow[ll, "h \mid B"] \arrow[uu, "g"'] \\
(M, M - x_j) & &
\end{tikzcd}$$

DIAGRAM (13)

We take $\mu^*_j = h^{*-1}(\eta)$ and observe that since ρ is a deformation retraction of E^n onto B, $\rho^* = s^{*-1}$. Therefore, since Diagram (13) commutes, $I_f(x_j) = i_f(x_j)$.

UNIVERSITY OF CALIFORNIA,
LOS ANGELES.

REFERENCES.

[1] F. Browder, "On the fixed point index for continuous mappings of locally connected spaces," *Summa Brasiliensis Mathematica*, vol. 4 (1960), pp. 253-293.

[2] M. Brown, "A mapping theorem for untriangulated manifolds," *The Topology of 3-Manifolds*, M. Fort (ed.), Prentice-Hall, Englewood Cliffs, New Jersey, 1962, pp. 92-94.

[3] R. Brown, "On the Lefschetz number and the Euler class," *Transactions of the American Mathematical Society*, to appear.

[4] E. Fadell, "Generalized normal bundles for locally flat imbeddings," *to appear*.

[5] H. Hopf, "Über die algebraische Anzahl von Fixpunkten," *Mathematische Zeitschrift*, vol. 29 (1929), pp. 493-524.

[6] S. Lefschetz, "Intersections and transformations of complexes and manifolds," *Transactions of the American Mathematical Society*, vol. 28 (1926), pp. 1-49.

[7] ———, "Manifolds with a boundary and their transformations," *Transactions of the American Mathematical Society*, vol. 29 (1927), pp. 429-462.

[8] J. Leray, "Sur les équations et les transformations," *Journal de Mathématiques Pures et appliquées*, 9 me serie, vol. 24 (1945), pp. 201-248.

[9] J. Weier, "Fixpunkttheorie in topologischen Mannigfaltigkeiten," *Mathematische Zeitschrift*, vol. 59 (1953), pp. 171-190.

THE RANK OF $S^2 \times S^1$.

By Harold Rosenberg.*

The *rank* of a differential manifold V is the maximal number of linearly independent vector fields on V which pairwise commute. In this paper we will prove the rank of $S^2 \times S^1$ is one. If V is a closed manifold the rank of V is the largest integer k such that there is an action of R^k on V all of whose orbits are k dimensional. These two numbers are not the same in general. For example, let $V = S^2 \times R$. The rank of V is three, but there is no action of R^3 on V all of whose orbits are three dimensional, since the orbit of a point would be open and closed in V hence all of V. It would be interesting to know if there is an action of R^3 on this space all of whose orbits are two dimensional.

E. Lima has recentily shown the rank of S^3 is one, [3]. Some of the ideas of this paper have their genesis in this work of Lima.

I wish to express my gratitude to my colleagues for the generous assistance I received in the preparation of this paper; in particular, conversation with A. Haefliger, E. Lima, R. Sacksteder, and S. Smale were particularly useful.

Definitions and Notation. We adopt the terminology of [3]. Let V^n be a closed manifold (everything of class C^∞), and $X_1, \cdots, X_k$ vector fields on V^n with integral curves $\xi^1, \cdots, \xi^k$ respectively. The vector fields *commute* if $[X_i, X_j] = 0$ for all i and j. In terms of integral curves this means $\xi_t^i \circ \xi_s^j = \xi_s^j \circ \xi_t^i$ for all real numbers s and t.

An *action* ϕ of a Lie Group G on V is a differentiable may $\phi\colon G \times V \to V$ such that (i) $\phi(gh, x) = \phi(g, \phi(h, x))$ for all $g, h \in G$ and $x \in V$, and (ii) $\phi(e, x) = x$ for $x \in V$, e the identity of G. For $x \in V$, the *isotropy subgroup* of x is $H_x = \{g \in G/\phi_g(x) = x\}$ (here we write $\phi(g, x) = \phi_g(x)$), it is a closed subgroup of G. The *orbit*, or *leaf* of x is $\{\phi_g(x)/g \in G\}$. The action ϕ induces by passing to the quotient space, a 1-1 continuous map of G/H_x onto L_x, the orbit of x.

Received February 20, 1964.

* This work was partially supported by the Office of Naval Research under Contract No. NONR(G)-00011-64-M.

When $G=R^k$ and V^n is closed, an action ϕ is equivalent to k commuting vector fields on V^n; the relation is

$$\phi(t,x)=(\xi_{t_1}^1\circ\xi_{t_2}^2\circ\cdots\circ\xi_{t_k}^k)(x),\ t=(t_1,\cdots,t_k)\in R^k.$$

We call ϕ a *locally free action* if all the orbits are k-dimensional.

Our main interest is $n=3$, $k=2$. The orbits of x are classified by their isotropy subgroups H_x and we have the following possibilities: the dimension of H_x is 2. Then $H_x=R^2$ and $L_x=x$. The dimension of H_x is one. Then $H_x=L+nv$, L a line through the origin, v a vector in R^2 and $n=0,\pm1,\pm2,\cdots$. L_x is then a line or circle (i.e., 1-1 continuous image of) depending on the direction of v. The case dimension $H_x=0$ gives three possible orbits. When $H_x=Zu$, Z the group of integers, $u\in R^2$, we have $L_x=R^2$ or a cylinder, depending on whether $u=0$ or $u\neq 0$. If $H_x=Zu+Zv$, u and v linearly independent, then L_x is a torus.

A subset A of V is *invariant* under the action ϕ if $x\in A$ implies $\phi(G,x)\subset A$. A minimal set of the action ϕ is a compact invariant non-empty set with no proper compact invariant subsets. If V is compact, every compact invariant set contains a minimal set.

We will write $A \,/\!\!/\, B$ to mean the intersection of A and B is empty.

1. The existence of non simply connected leaves. In this section we use recent results of R. Sacksteder to find conditions on actions which guarantee the existence of non simply connected leaves. For the applications to this paper we need to know every locally free action of R^2 on $S^2\times S^1$ has a non simply connected leaf. This already follows from Haefliger's thesis [2], without using the results of Sacksteder. However, the theorem we obtain is quite general. We assume the reader is familiar with the definition of foliation and holonomy, but this is not necessary for § 2 and § 3. For definitions of these notions we refer the reader to [5].

THEOREM 1.1. (Sacksteder [5]). *Let V be a closed manifold together with a foliation of codimension one such that all the leaves have finite holonomy groups. Then there is a Riemannian metric on V which is invariant under holonomy.*

We make this precise below. Let x and y be points of the same leaf L and C be a curve in L joining x to y. For each transversal arc A at x and B at y the holonomy defines a diffeomorphism of a neighborhood of x on A to a neighborhood of y on B. Thtorem 1.1 asserts: there is a Riemannian metric on V such that if A and B are orthogonal trajectories at x and y

respectively, then the induced local diffeomorphism is an isometry. Intuitively, the distance between two leaves remains constant.

THEOREM 1.2. *Let V be a closed n manifold and ϕ an action of R^{n-1} on V such that isotropy subgroup of each point is zero. Then there is a covering map $p: R^n \to V$.*

Proof. The action ϕ induces a foliation of codimension one and the isotropy subgroup of each point being zero means the orbit of each point is the one to one continuous image of R^{n-1}. Thus each orbit is simply connected hence has a trivial holonomy group and Theorem 1.1. applies. Assume then, that V has a metric invariant under holonomy.

Let $x \in V$ and $f: R \to V$ be a parametrization of the orthogonal trajectory through x such that $t =$ the length of the trajectory between x and $f(t)$, that is,

$$t = \int_0^t \| f'(s) \| \, ds.$$

Such a parametrization exists because V is compact hence complete and the orthogonal trajectories are geodesics hence infinitely extendable.

Define $p: R^n \to V$ by

$$p(r,t) = \phi_r(f(t)), \qquad (r,t) \in R^{n-1} \times R$$

We will show p is a covering map.

Let ω be the one form on V defined as follows. At a point $x \in V$ choose a coordinate system $x_1, \cdots, x_n$, such that $x_1, \cdots, x_{n-1}$ are coordinates of a neighborhood of x in L_x and x_n is the parametrization by arc length of the orthogonal trajectory to L_x at x. Define ω at p to be dx_n. Since the foliation is invariant under holonomy, ω is a closed one form on V such that

$$\int_h \omega = 0,$$

if h is a curve entirely contained in one leaf, and

$$\int_h \omega = \text{length of } h,$$

if h is a segment of an orthogonal trajectory. We assume V is oriented when we speak of the orthogonal trajectory, but this is no loss of generalization since we may pass to the oriented two sheeted covering of V and lift the action to this space.

Let H be the subgroup of $\pi_1(V, x)$ consisting of homotopy classes which may be represented by loops h at x such that

$$\int_h \omega = 0.$$

Since ω is a closed form, H is a well defined subgroup. Let W be the covering space of V defined by H; W is the space of equivalence classes of maps $h: I \to V$ with $h(0) = x$ under the relation $h_1 \sim h_2$ if $h_1(1) = h_2(1)$ and

$$\int_{h_1} \omega = \int_{h_2} \omega$$

The covering map $g: W \to V$ is $g(h) = h(1)$. Let ϕ_1 be the action of R^{n-1} on W obtained by lifting ϕ, so that $g\phi_1 = \phi(1 \times g)$, $1 =$ the identity map of R^{n-1}.

For each real number t, define a map $U(t): I \to V$ by $U(t)(s) = f(ts)$. Let $k: R \to W$ be the map $k(t) = (U(t)) =$ the equivalence class of $U(t)$. Since

$$\int_{U(t)} \omega = t,$$

k is an injection.

Define $j: R^n \to W$ by $j(r, t) = \phi_1(r, k(t))$, $(r, t) \in R^{n-1} \times R$. Since $gj = p$, p is a covering map if j is a diffeomorphism. The orbits of ϕ_1 are R^{n-1} hence j is a diffeomorphism if each orbit of ϕ_1 intersects $k(R) = S$ in exactly one point.

First we will show that each leaf meets S in at most one point. Let the leaf A of ϕ_1 meet S at (h_1) and (h_2). This means $h_1(1) = f(t_1)$, $h_2(1) = f(t_2)$ for some t_1, t_2, $h_1(0) = h_2(0) = x$, and

$$\int_{h_1} \omega = t_1; \qquad \int_{h_2} \omega = t_2.$$

Since (h_1) and (h_2) are in the same leaf A there is a curve in A joining the two points; that is, there is a map $F: I \times I \to V$ with $F(0, t) = h_1(t)$, $F(1, t) = h_2(t)$, $F(s, 0) = x$ and $F(s, 1)$ is in the leaf containing $h_1(1)$ for $0 \leqq s \leqq 1$. Let $h_3(t) = F(t, 1)$ for $t \in I$. F is a homotopy, rel. $\{0, 1\}$, between the curves $h_3 \circ h_1$ and h_2, hence

$$\int_{h_3 \circ h_1} \omega = \int_{h_2} \omega$$

Now observe that h_3 is a curve contained in one leaf, and

$$\int_{h_2} \omega = t_2,$$

hence

$$\int_{h_3 \circ h_1} \omega = \int_{h_3} \omega + \int_{h_1} \omega = \int_{h_1} \omega = t_1.$$

Thus $t_1 = t_2$ and $(h_1) = (h_2)$ hence each orbit of ϕ_1 meets S in at most one point.

Now let (h) be a point of W and A be the orbit of ϕ_1 through (h). We will prove A meets S by constructing a map $G: I \times I \to V$ satisfying: $G(1,t) = h(t)$, $G(0,t) = U_a(t)$ for some real number a, $G(s,0) = x$ and $G(s,1)$ is in the leaf through $h(1)$ for $0 \leqq s \leqq 1$. The map $s \to (G(s,\ ))$ is then a curve in W joining (h) to (U_a). By $G(s,\ )$ we mean the map $G(s,\ )(t) = G(s,t)$. Since (U_a) is a point of S this will complete the proof. Observe that a curve h in V is homotopic to a curve consisting of segments such that each segment is an arc of an orthogonal trajectory or is contained in one leaf. Therefore we may assume there exist numbers $0 = t_0 < t_1 < \cdots < t_k = 1$ such that for each i, the arc $h[t_i, t_{i+1}]$ is either a segment of an orthogonal trajectory or is contained in one leaf.

Let L be the leaf containing x, $C(t)$ be a curve in L starting at x, and s_0 a positive real number. Since the orthogonal trajectories are infinitely extendable, the orthogonal trajectories of length s_0 along C define a map $F: I \times [0, s_0] \to V$ such that, for fixed t, $F(t,s)$ is an orthogonal trajectory with $F(t,0) = C(t)$, and $F(t,s)$ is the point a distance s from $C(t)$ along the orthogonal trajectory. Moreover, the metric on V is invariant under holonomy, hence the points $F(t,s)$ for fixed s are contained in the leaf through $f(s)$.

Now G is defined as follows. We may assume $h[t_0, t_1]$ is contained in L and $h[t_1, t_2]$ is an orthogonal trajectory. Let C be the curve $h[t_0, t_1]$ and s_0 the length of $h[t_1, t_2]$. Apply the last paragraph to obtain a map $F_1: I \times [0, s_0] \to V$ such that $F_1(0,s) = f(s)$, $F_1(1,s) = h(t_1 + s)$ and $F_1(t, s_0)$ is in the leaf through $f(s_0)$ for $0 \leqq t \leqq 1$. Repeat this with C the curve $F_1(t, s_0)$ followed by $h(t_2, t_3)$. Induction on k yields the desired map G. This completes the proof of Theorem 1.2.

Remark. It has been suggested to me by André Haefliger that if V satisfies the hypothesis of Theorem 1.2 then V may be covered by $R^{n-1} \times S^1$ as follows. Let $S = S^1 \to V$ be a diffeomorphism of S^1 into V such that $F(S^1)$ is orthogonal to the foliation determined by ϕ (one can always find closed orthogonal trajectories, [5]). Let $p: R^{n-1} \times S^1 \to V$ be the map $p(r,t) = \phi_r(F(t))$, then p is a covering map. Using Theorem 1.1 Haefliger has proved this fact (unpublished). The proof we have given of Theorem 1.2

will also show p is a covering map. One must replace the subgroup H of $\pi_1(V,x)$, $H=\{[\alpha]; \int_\alpha \omega=0\}$, by the subgroup H_1 generated by FS^1. Then the covering space of V determined by H_1 is seen to be $R^{n-1}\times S^1$.

COROLLARY 1.3. *Let V be a closed n-dimensional manifold which can not be covered by $R^{n-1}\times S^1$. Every locally free action of R^{n-1} on V has a non-simply connected leaf.*

2. Preliminaries. In this section we derive some topological properties of $S^2\times S^1$. We consider $S^2\times S^1$ as the quotient space of $S^2\times I$ under the identification $(x,0)=(x,1)$, $x\in S^2$, and will denote this space by V. Let $S=S^2\times\frac{1}{2}\subset \mathrm{V}$.

LEMMA 2.1. *Let W be an embedded sphere in V. Then W bounds a ball in V or W represents the non zero element of $H_2(V,Z_2)$.*

Proof. First we will prove 2.1 under the additional hypothesis that W is disjoint from S. For each real number t let $[t]$ be the greatest integer less than or equal to t. The map $p: S^2\times R\rightarrow V$ defined by $p(x,t)=(x,t-[t])$ is a covering map. Since W is simply connected there is a sphere E in $S^2\times R$ such that $p(E)=W$. We have $E\subset S^2\times(k-\frac{1}{2},k+\frac{1}{2})$ for some integer k. $S^2\times R$ is diffeomorphic to R^3-0 hence by Schoenflies theorem, E is the boundary of a ball F in $S^2\times R$ or $E=\partial(F-0)$. In the latter case, E represents the generator of $H_2(S^2\times R,Z_2)$ and $p_\#: H_2(S^2\times \mathrm{R})\rightarrow H_2(V)$ is an isomorphism, hence $p\cdot(E)\doteq W$ represent the generator of $H_2(V,Z_2)$. If $E=\partial F$ then $F\subset S^2\times(k-\frac{1}{2},k+\frac{1}{2})$ and p is a diffeomorphism on $S^2\times(k-\frac{1}{2},k+\frac{1}{2})$, hence W bounds the ball $p(F)$. Clearly S could be any sphere $S^2\times t$, $t\in I$, hence 2.1 is true for any sphere W which is disjoint from some sphere $S^2\times t$.

To complete the proof it suffices to show: if W is an embedded sphere in V, then W is isotopic to a sphere which is disjoint from S. For, according to [1], if A and B are isotopic submanifolds of V, there is a diffeomorphism of V which takes A onto B.

Suppose W is a sphere in V, not necessarily disjoint from S. The Thom Transversality Theorem implies there is a sphere in V transverse to S and isotopic to W. Hence we may assume $W\cap S=\beta_1\cup\cdots\cup\beta_n$, each β_i a simple closed curve. Choose β_i so that one of the connected components of $W-\beta_i$ contains no β_j in it's interior. Denote this component by D. Let M and N be the connected components of $S-\beta_i$. Since $D\cap S=\beta_i$, there is a t such that the sphere $D\cup M$ is disjoint from $S^2\times t$. Thus $D\cup M$

bounds a ball or represents the non zero element of $H_2(V)$. Similarly for the sphere $D \cup N$. If both of these spheres represented the non zero element of $H_2(V)$, then their sum in this group would be zero. However, with $+$ denoting addition in $H_2(V)$, we have:

$$(D \cup M) + (D \cup N) = 2D + (M \cup N) = M \cup N = S.$$

Hence one of the spheres bounds a ball in V; assume $D \cup M = \partial B$, B a ball.

Let U be a closed tubular neighborhood of B; U is diffeomorphic to $B \times I$. Choose corrdinates x, y, z, of U satisfying:

$$\begin{aligned} U &= \{(x,y,z)\,;\ x^2+y^2+z^2 \leqq 1+\epsilon,\ \epsilon > 0\} \\ B &= \{(x,y,z)\,;\ x^2+y^2+z^2 \leqq 1,\ z \geqq 0\} \\ D &= \{(x,y,z)\,;\ x^2+y^2+z^2 = 1,\ z \geqq 0\} \\ &\quad \{(x,y,z)\,;\ 1 \leqq x^2+y^2 \leqq 1+\epsilon,\ z=0\} \subset S. \end{aligned}$$

Choose ϵ sufficiently small so that

$$\{(x,y,z)\,;\ 1 < x^2+y^2+z^2 \leqq 1+\epsilon,\ z \geqq 0\}$$

is disjoint from W. Let E be the disk

$$\{(x,y,z)\,;\ x^2+y^2+z^2 = 1+\epsilon/2,\ z \geqq 0\}.$$

The boundary of E bounds the disk

$$F = \{(x,y,z)\,;\ x^2+y^2 \leqq 1+\epsilon/2,\ z=0\}$$

which is contained in S. $E \cup F$ bounds the ball

$$\{(x,y,z)\,;\ x^2+y^2+z^2 \leqq 1+\epsilon/2,\ z \geqq 0\}$$

hence E is isotopic to F by an isotopy leaving ∂E fixed. The map $F \times I \to V$, $((x,y,0),t) \to (x,y,(1-t)\,((1+\epsilon/2)^2 - x^2 - y^2)^{\frac{1}{2}})$ is an isotopy with these properties.

Let R be the sphere $E \cup (S - F)$. R is isotopic to S; the isotopy is defined to be the identity on $S - F$ together with the above isotopy between E and F. By definition of R we have, $R \cap W = \beta_1 \cup \cdots \cup \hat{\beta}_i \cup \cdots \cup \beta_n$. Thus W is isotopic to a sphere meeting S in $n-1$ simple closed curves. Induction on n completes the proof.

LEMMA 2.2. *Let A be a closed 2-dimensional manifold imbedded in V. There is a diffeomorphism f of V such that $f(A) \cap S$ contains no simple closed curves which are null homotopic on A.*

Proof. We may assume A intersects S transversally, so that $A \cap S = \beta_1 \cup \cdots \cup \beta_n$, each β_i a simple closed curve. Let $\beta_1 \cdots, \beta_k$, be the curves which are null homotopic on A; each such β_i bounds a disk D_i contained in A. Choose i so that D_i contains no β_j in it's interior. Let M and N be the connected components of $S - \beta_i$. According to 2.1, the sphere $D_i \cup M$ bounds a ball or represents the non zero element of $H_2(V, Z_2)$. Similarly for the sphere $D_i \cup N$. Exactly as in the proof of 2.1, one proves that both spheres can not be non zero in this group; hence we may assume $D_i \cup M$ bounds a ball B. Now proceed as in 2.1; displace D_i across the ball B to obtain an isotopy between A and a manifold whose intersection with S is $\beta_1 \cup \cdots \cup \hat{\beta}_i \cup \cdots \cup \beta_n$. Induction on k yields the desired diffeomorphism.

LEMMA 2.3. *Let T be a torus topologically embedded in V such that $T \cap S = \phi$. Then T or $T + S$ separates V into two components of which it is the complete point set boundary.*

Proof. Recall that if M^{n-1} is a connected submanifold of N^n and $i_\#: H_{n-1}(M^{n-1}; Z_2) \to H_{n-1}(N^n; Z_2)$ is zero, $i_\#$ the map induced by the inclusion $i: M^{n-1} \subset N^n$, then M^{n-1} separates N^n into 2 components of which it is the complete set boundary [3]. If $i_\#(T) = 0$, then T bounds and we are done; if $i_\#(T) \neq 0$ then $i_\#(T + S) = i_\#(T) + i_\#(S) = 0$ in $H_2(V; Z_2)$, hence $T + S$ bounds, and it is necessarily the complete boundary. By considering the exact sequence of the pair $(V, T + S)$, it is easy to see $T + S$ separates V into 2 components.

LEMMA 2.4. *Let β be a simple closed curve on S, e_1, e_2 a field of 2 frames on β such that e_1 is tangent to β and e_2 is transverse to S. Then e_1, e_2 cannot be extended to a field of 2 frames on V; indeed, there is no continuous extension of e_1, e_2 to either component of $S - \beta$.*

Proof. This follows easily from the following observation of Reinhart [4]: let $S^1 = \{(x, y)/x^2 + y^2 = 1\} \subset R^2$, e_1, e_2 be the 2 frame in R^3 on S^1 defined by $e_1(x, y) = (0, 0, 1)$, $e_2(x, y) = (-y, x, 0)$, then e_1, e_2 cannot be extended to a continuous 2 frame on $D^2 = \{(x, y)/x^2 + y^2 \leqq 1\}$. The obstruction to extending this 2 frame represents the non-zero element of $\pi_1(SO(3))$.

3. Compact orbits. In this section we prove theorems leading to the existence of a compact leaf of a locally free action on $V = S^2 \times S^1$. Let ϕ be any action of R^2 on V, X and Y the commuting vector fields induced by ϕ.

THEOREM 3.1. *Let μ be a cylindrical orbit of ϕ whose closure is disjoint from S. Then ϕ has a toral or circle orbit.*

Proof. The topological torus constructed in [3] is the main tool of the proof. We will indicate how it is done but the details are extensive so we refer the reader to [3] for complete proofs.

Let $x_0 \in \mu$. The isotropy subgroup H_{x_0} is a discrete group on one generator so by a judicious choice of a, b, c, d we may be sure the vector fields $X' = aX + bY$, $Y' = cX + dY$ have μ as a cylindrical orbit, commute, and the X' orbit of X_0 is closed. Take $X' = X$, $Y' = Y$ and the X orbit of X_0 closed.

Since X and Y commute, the X orbit of each $x \in \mu$ is closed and of the same period. Let K be a minimal set in the closure of μ. Then $K \not\parallel S$ and the X orbit of each $x \in K$ is closed and of the same period; since two points of a minimal set have the same isotropy subgroup.

Let $x \in K$. If $X(x)$ and $Y(x)$ are linearly dependent, the ϕ orbit of x coincides with the X orbit of x and is a circle. If $X(x)$ and $Y(x)$ are independent, the ϕ orbit is a cylinder or torus. To complete the proof it suffices to show the orbit cannot be a cylinder.

Assume L_x is a cylinder. Let γ_1 be the closed X orbit through x and S_0 be a cylindrical band about γ_1, spanned by geodesics normal to the ϕ orbit through γ_1; i.e., normal to L_x. Since K is a minimal set of ϕ, given any neighborhood U of γ_1, there are arbitrarily large values of t such that $\eta_t(\gamma_1) \subset U$, η_t is the integral curve of the Y vector field. Define a real valued function $\tau(x)$ on γ_1 by $\tau(x) =$ the smallest positive number such that $\eta_{t(x)}(x) \in S_0$. As shown in [3], we may choose S_0 so that $\tau(x)$ becomes a differentiable map inducing a diffeomorphism of γ_1 onto $\gamma_2 \subset S$ defined by $x \to \eta_{t(x)}(x)$ and satisfying: there is a ring $A \subset S_0$, A is diffeomorphic to $S^1 \times I$ and A has boundary $\gamma_1 + \gamma_2$.

Let $B = \{\eta_t(x)/x \in \gamma_1, 0 \leqq t \leqq \tau(x)\}$ and $T = A \cup B$; T is a topological torus in V. Since $K \not\parallel S$ we may choose S_0 so small that $S_0 \not\parallel S$ hence $T \not\parallel S$.

According to 2.3, either T separates V or $T + S$ separates V. Let $V - T = C_1 \cup C_2$, C_1, C_2 connected and disjoint with $\partial C_1 = T = \partial C_2$ [Fig. 1]. For small ϵ, the set $\{\eta_t(x)/x \in A - B, 0 < t \leqq \epsilon\}$ lies entirely in one component, C_1 say; and $\{\eta_t(x)/x \in A - B, -t \leqq \epsilon < 0\}$ lies in C_2. Also $\{\eta_t(y)/y \in \gamma_1, -\epsilon \leqq t < 0\}$ is contained in C_2 and $\{\eta_t(y)/y \in \gamma_2, 0 < t \leqq \epsilon\}$ is contained in C_1.

Let $\gamma = \eta_{-\epsilon}(\gamma_1) \subset C_2$, and $w \in \gamma$. Because K is a minimal set there are arbitrarily large t such that $\eta_t(w)$ is as close to γ as we wish; hence $\eta_t(w) \in C_2$ for arbitrarily large t. But this is impossible since $\eta_t(w) \in C_1$ for $t \geqq \tau(w) + \epsilon$; i.e., $\eta_t(w) \in C_1$ for $\tau(w) < t < \tau(w) + \epsilon$ and as t increase $\eta_t(w)$ may enter C_2 only by crossing A or B. It can't cross A because the Y vector field is

normal to A pointing into C_1 and we can't cross B because it is a cylindrical orbit. Hence if T separates V we know there is a toral orbit.

Now suppose $T+S$ separates V; $V-(T+S)=C_1+C_2$, $C_1 /\!/ C_2$, $\partial C_1 = T+S = \partial C_2$, $T+S$ the complete point set boundary. We use the same notation as above. Let $\gamma^1 = \eta_{+\epsilon}(\gamma_2)$. By lifting the action to $S^2 \times R$, it is simple to check: γ^1 is homotopic to zero in C_1 iff γ is not homotopic to zero in C_2.

Assume then that γ is not homotopic to zero in C_2 (the case $\gamma^1 \not\simeq 0$ is

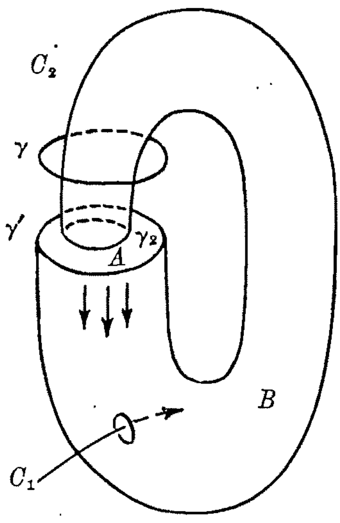


FIG. 1.

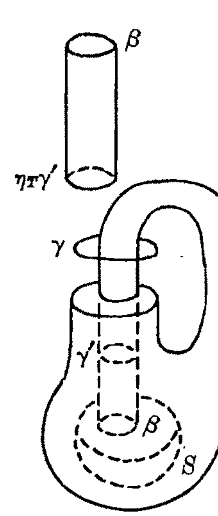


FIG. 2.

treated similarly). By the minimality of K, there are arbitrarily large values T such that $\eta_T\gamma^1$ is so close to γ that $\eta_T\gamma^1 \subset C_2{}^0$ and $\eta_T\gamma^1 \not\simeq 0$ in C_2 [Fig. 2].

Since $\eta_t(\gamma^1)$ is disjoint from T for $t \geqq \epsilon$, the cylinder $L = \{\eta_t(\gamma^1)/\epsilon \leqq t \leqq T\}$ must intersect S. We may assume this intersection is transverse. By Lemma 2.2 we may suppose each of the components of $L \cap S$ is a generator of $\pi_1(L)$. Let β be the component of $L \cap S$ such that the cylinder on L bounded by β and $\eta_T(\gamma^1)$ lies entirely in C_2. Now β bounds a disk on S and β is isotopic to $\pi_T(\gamma^1)$ in C_2 so $\eta_T(\gamma^1)$ also bounds a disk in C_2. But this contradicts $\pi_T(\gamma^1) \not\simeq 0$ in C_2 and completes the proof of 3.1.

THEOREM 3.2. *Let μ be a cylindrical orbit of ϕ contained in a minimal set. If some generator C of $\pi_1(\mu)$ is disjoint from S, then ϕ has an orbit of dimension less than two.*

Proof. Choose $x_0 \in \mu$ so the X orbit Γ through x_0 is disjoint from C.

Let τ be so large that the cylinder $L = \{\eta_t C / -\tau \leqq t \leqq \tau\}$ contains C and Γ in its interior. L is isotopic to a cylinder disjoint from S so we may assume L is disjoint from S. Let g be a diffeomorphism of L onto L, isotopic to the identity of L and $g\Gamma = C$. Then g extends to a diffeomorphism $\tilde{g}$ of V. Considering the action induced by $\tilde{g}$ shows we may consider C to be a closed X orbit.

Now we construct the topological torus T as in 3.1, making sure $A \cap S = \phi$. We can do this because $C /\!\!/ S$. T is not a differentiable submanifold of V but the corners on T, in the notation of 3.1, the curves γ_1, γ_2, do not meet S. Hence we may assume T intersects S transversally.

If T is disjoint from S, Theorem 3.1, Figure 2, implies T is a toral orbit, contradicting our hypothesis. Let $T \cap S = \beta_1 \cup \cdots \cup \beta_n$, the β_i simple closed curves, pairwise disjoint. According to 2.2 we may suppose no β_i is null homotopic on T.

Let β_1 be a generator of $\pi_1(T)$. Since $\beta_1 \cap C = \phi$, β_1 is isotopic to C on $(T - A^0)$. This allows us to construct a map f of μ which is the identity outside a compact cylinder E which contains β_1 and C, such that f/E is isotopic to the identity of E and $f(\beta_1) = C$. Then f/E extends to a diffeomorphism $\bar{f}$ of V and considering the action induced by $\bar{f}$ we can take $C = \beta_1$, μ transverses to S along β_1.

Let $e_1 = X/C$, $e_2 = Y/C$. By Lemma 2.4, e_1, e_2 cannot be extended to a 2-frame on either component of $S - C$. Hence at some point of S, X and Y are colinear. Then the orbit of ϕ through this point has dimension less than two.

The next proposition is easy to prove; we leave this to the reader.

LEMMA 3.3. *Let $z \in S$ and x^1, x^2, x^3 be a coordinate system in a chart U about z. There is a sphere $\bar{S}$ in V, isotopic to S, and a neighborhood $U^1 \subset U$ such that $\bar{S} \cap U^1 = \{(x^1, x^2, x^3)/x^2 = 0\}$.*

THEOREM 3.4. *Let μ be a cylindrical orbit of ϕ contained in a minimal set. There is a diffeomorphism f of V such that some generator of $\pi_1(f(\mu))$ is disjoint from S.*

Proof. Let C be a generator of $\pi_1(\mu)$ which, as in the proof of 3.2, is also a closed orbit of the X vector field. We may assume $C \cap S \neq \phi$ and S is transverse to μ in a neighborhood of C.

Let $C \cap S = \{z_1, \cdots, z_n\}$. It is known that for each z_i there is a chart U_i with coordinates $x_1{}^i$, $x_2{}^i$, $x_3{}^i$, such that $z_i = (0,0,0)$ and $X/U_i = \partial/\partial x_2{}^i$, $Y/U_i = \partial/\partial x_1{}^i$. Choose the U_i so small that $U_i /\!\!/ U_j$ if $i \neq j$. According to 3.3, we may assume $U_i \cap S =$ the plane $x_2{}^i = 0$. In U_i, $C = \{(0, x_2{}^i, 0)\}$.

Let $F_i = \{(x_1{}^i, 0, x_3{}^i)/-1 \leqq x_j{}^i \leqq 1, j = 1, 3\} \subset U_i \cap S$. Since μ is

252560

contained in a minimal set and $\eta_t C$ is a closed X orbit for all t, we can choose $t_1 > 0$ so that $C_1 = \eta_{t_1} C$ intersects S in n points, one point in each F_i. Let $x_3{}^i(C_1)$ be the $x_3{}^i$ coordinate of the point of intersection of C_1 with F_i. Since μ is not a torus we have $x_3{}^i(C_1) \neq 0$. Choose $t_2 > t_1 + 2$ so that $\eta_{t_2} C$ meets S in n points, one in each F_i, and $|x_3{}^i(C_2)| < |x_3{}^i(C_1)|$ for $1 \leqq i \leqq n$. Let $t_j > t_{j-1} + 2$ satisfy: $\eta t_j C$ meets S in n points, one in each F_i, and $|x_3{}^i(C_j)| < |x_3{}^i(C_{j-1})|$ for $1 \leqq i \leqq n$ and $2 \leqq j \leqq N$. We will choose N later.

Let $B = \{\eta_t C/0 \leqq t \leqq t_N\}$. Define $H_i = \{(x_1{}^i, x_2{}^i, x_3{}^i\} | x_j| \leqq 1, j = 1, 2, 3/$ and $H = \bigcup_{i \leqq n} H_i$. B meets H transversally hence there is a neighborhood N_0 of H and a compact cylinder $B^1 \supset B \cap H = B^1 \cap H$, B^1 meets S transversally, and B^1 is as close to B as we wish. Then there is a diffeomorphism of V which is the identity on H and takes B onto B^1. Therefore we may assume B meets S transversally.

$B \cap S$ is a closed one dimensional submanifold of B. The components of $B \cap S$ may be (i) simple closed curves contained in the interior of B, (ii) homeomorphs of $[0, 1]$ with both end points in the same component of ∂B, or (iii) homeomorphs of $[0, 1]$ with one end point in C, the other in $\eta_{t_N} C$.

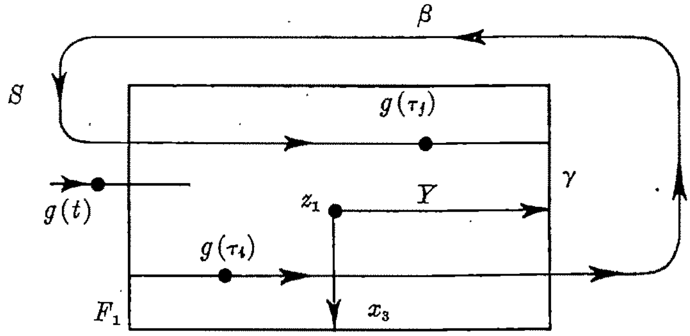
We now prove there are no components of type (iii). Suppose $g: I \to B \cap S$ is of type (iii), $g(0) \in \eta_{t_N} C$ and g is 1-1. Due to the representation of X and Y in U_i, if g meets F_i, the intersection is a curve $x_3{}^i =$ constant, $x_2{}^i = 0, -1 \leqq x_1{}^i \leqq 1$.

It is important to observe the following property of g: Suppose $g(t_1) \in F_i$, $g(t_2) \notin F_i$, $t_1 < t_2$; let t be the smallest number greater that t_2 such that $g(t) \in F_i$ (if there is such a number), then $g(t) = (-1, 0, x_3)$ for some x_3, $-1 \leqq x_3 \leqq +1$. Otherwise $g(t) = (1, 0, x_3)$ and the two frames $\{X(g(t_1)), Y(g(t_1))\}$, $\{X(g(t)), Y(g(t))\}$ define different orientations of B.

Assume N was chosen larger than $n + 1$. The image of g meets each C_i hence there exist $\tau_1 < \tau_2 < \cdots < \tau_{n+1}$ such that $g(\tau_i) \in F_{K_i}$ for some K_i. Since there are only n F_j's and $n + 1$ distinct τ's, there are $\tau_i, \tau_j \ni \tau_i < \tau_j$ and $g(\tau_i)$, $g(\tau_j)$ are in the same F, F_1 say. Also, because $t_i > t_{i-1} + 2$ for each i, we know $g(t)$ for $\tau_i \leqq t \leqq \tau_j$ leaves F_1 before reentering F_1 at τ_j. This implies that for $k > j$, $g(I) \cap C_k$ is not in F_1; i.e., $g(t)$ does not reenter F_1 for $t > t_j + 2$ at any C_k. One can see this by observing that for $t > t_j + 2$ the only way $g(t)$ can enter F_1 is by crossing the curve $x_1{}^1 = -1$, $x_2{}^1 = 0$, $-1 \leqq x_3{}^1 \leqq 1$ and at this crossing the absolute value of its x_3 coordinate must be less than the absolute value of the x_3 coordinates of $g(\tau_i)$ and $g(\tau_j)$. Let $\tau_i{}^1$ be the smallest t larger than τ_i such that $g(\tau_i{}^1) \in F_1$ and $\tau_j{}^1$ be the smallest t larger than τ_j such that $g(\tau_j{}^1) \in F_1$. Consider the curve

$$\beta = \gamma \cup \{g(t)/\tau_i{}^1 \leqq t \leqq \tau_j{}^1\}$$

where $\gamma = \{(1, 0, x_2)/x_3(g(\tau_i^1)) \leqq x_3 \leqq x_3(g(\tau_j^1))\}$. This is a simple closed curve on the sphere S and in order for g to reenter F_1 as described, g would have to intersect β at $\{g(t)/\tau_i^1 \leqq t \leqq \tau_j^1\}$, contradicting our assumption that g has no self-intersections.



Now assume $N > 2n + 2$. The image of g meets $C_{n+2}, \cdots, C_{2n+2}$ and at each intersection g is in some F_j for $j \neq 1$. There are n distinct points on g at which g is in some F_j, $j \neq 1$ and there are $n - 1$ F_j's g may be in, so there is some F_j, $j = 2$ say, such that g enters F_2, then leaves with increasing time, and reenters F_2 with a smaller x_3 coordinates. Exactly as before we see that g may not enter F_2 when g meets C_j for $j > 2n + 2$. If we choose N large enough so that we can continue this process $n - 2$ more times, we conclude that for large j, $g \cap C_j$ cannot be in any F_i for $1 \leqq i \leqq n$. This is clearly impossible. This proves no curves of type (iii) exist in $B \cap S$.

According to Lemma 2.2, we may assume no curves of type 1 exist. Now let C^1 be a simple closed curve on B which is a generator of $\pi_1(B)$ and disjoint from S. C^1 exists because $B \cap S$ is of type (ii). Let f be a diffeomorphism of μ with compact support, isotopic to the identity of μ and taking C onto C^1. Then f extends to a diffeomorphism of V. This completes the proof of 3.4.

COROLLARY 3.5. *Let ϕ be a locally free action of R^2 on V, then ϕ has a compact leaf (a torus).*

Proof. According to Corollary 1.3, ϕ has a non simply connected leaf A; A is the one to one continuous image of a cylinder or torus. Let μ be an orbit contained in a minimal set in the closure of A. We have seen in the proof of 3.1, that μ must by a cylinder or torus. If it were a cylinder, then 3.2 and 3.4 imply ϕ has an orbit of dimension less than two hence is not locally free. Thus μ is a toral orbit of ϕ.

4. The main theorem.

THEOREM 4.1. *Let X and Y be commuting vector fields on $S^2 \times S^1$. There is some point at which X and Y are linearly dependent.*

Proof. We must show there is no locally free action of R^2 on $V = S^2 \times S^1$. Assume there is such an action. It follows from 3.5 that the action has a toral leaf T. We may assume T meets S transversally.

Suppose $T \cap S = \phi$. It is proved in [3] that if T is the boundary of a compact 3-manifold N such that the inclusion $i: T \subset N$ induces a map $i_*: \pi_1(T) \to \pi_1(N)$ with a non-zero kernel, then no pair of independent commuting vector fields on T can be extended to a two-frame on N. The proof in [3] also works if T is a component of δN. Now, according to 2.3, either T separates V or $T + S$ separates V. In any case, it follows from Van Kampen's theorem, that there is a component N of $V - T$ (or $V - (T + S)$) such that $i_*: \pi_1(T) \to \pi_1(N)$ has a non-zero kernel. Hence X and Y are dependent at some point of N.

Let $T \cap S = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_n$, where, by 2.2, each β_i is a generator of $\pi_1(T)$ and the β_i are pairwise disjoint simple closed curves. It is proved in [3] that there are real numbers a, b, c, d, such that the vector fields $X^1 = aX + bY$, $Y^1 = cX + dY$, are independent at each point and all the orbits of X^1 and Y^1 on T are closed. There is a diffeomorphism of T which is isotopic to the identity on T and takes an integral curve of X^1 or Y^1 onto β_1. Hence we may assume β_1 is an integral curve of X^1.

Consider the two frame S^1/β_1, Y^1/β_1. According to 1.4, this two frame does not extend to a two frame on S. This means ϕ is not a locally free action.

COROLLARY 4.2. *The rank of $S^2 \times S^1$ is one.*

COLUMBIA UNIVERSITY.

REFERENCES.

[1] J. Cerf, "Topologie de certains espaces de plongement," *Bulletin de la Société Mathématique de France*, vol. 89 (1961), pp. 227-380.

[2] A. Haefliger, *Varietes Feuilletees*. Annali della Scuola Normale Sup. di Pisa, Series III, vol. XVI, fasc. IV (1962), pp. 367-397.

[3] E. Lima, "Commuting vector fields on simply connected 3-manifolds," *Annals of Mathematics* (to appear).

[4] B. Reinhardt, "Periodic orbits on two manifolds," *Boletin de la Sociedad Matematica Mexicana*, vol. 5 (Oct. 1960), pp. 185-188.

[5] R. Sacksteder, "Foliations and Pseudogroups," *American Journal of Mathematics*, vol. 87 (1965), pp. 79-102.

IRREDUCIBLE SUBVARIETIES AND RATIONAL POINTS.

By NEWCOMB GREENLEAF.[1]

Suppose that f is a form (homogeneous polynomial) of degree d in n variables with coefficients in a field k. There is a considerable body of theorems to the effect that, when n is "large" with respect to d, then the form should have a non-trivial zero in k. In his address [4], partially devoted to such questions, Lang made the following conjecture. "Finally, to go back to number fields, it seems to me reasonable to expect that a form with $n > d$ at least has a non-trivial zero in all but a finite number of $\mathfrak{p}$-adic fields."

In this paper we prove this conjecture by elementary considerations of algebraic geometry. Given an algebraic set V defined over a field k (always assumed to be perfect), we consider the set of all subvarieties (absolutely irreducible) of V which are defined over k. Investigating this entity over various fields, and with respect to reduction $(\operatorname{mod} p)$, we arrive at our principal result, Theorem 2. Lang's conjecture is derived from this by a fairly standard argument.

Unless otherwise specified, all algebraic sets considered are affine. By a variety we shall exclusively mean an absolutely irreducible variety, by a k-variety an algebraic set, irreducible over k. When k is an algebraic number field, we shall consider k and all of its p-adic completions to be embedded in some fixed manner in the universal domain. Throughout we shall assume, without further reference, that all fields under consideration are perfect. The complications caused by inseparable extensions are indicated in Section 6.

An earlier version of this work formed a part of the author's Ph. D. thesis (Princeton, 1961). I should like to thank my supervisor, O. T. O'Meara, as well as Birch, Lewis, and Lang, for their advice and encouragement.

1. Irreducible subvarieties. If V is an algebraic set defined over k, we denote by $A(V,k)$ the union of all (absolutely irreducible) subvarities of V which are defined over k. For our purposes it will be much more useful

Received November 22, 1963.
Revised August 6, 1964.
[1] Work on this paper was partially supported by the National Science Foundation and the U. S. Army Research Office (Durham).

to regard $A(V,k)$ as given by the following construction. Set $V = \cup V_i = \cup V_{ij}$, where the V_i are the component k-varieties of V, and the V_{ij} are the absolute components of V_i. Recall that the V_{ij}, for fixed i, are conjugate varieties over k, under the automorphisms of K_i over k, where K_i is th smallest common field of definition of the V_{ij} over k (see [3, p. 73]). The algebraic set $W_i = \cap V_{ij}$, which is left fixed by all automorphisms of K_i over k, is defined over k. We set $A^1(V,k) = \cup W_i$. Setting $A^t(V,k) = A^1(A^{t-1}(V,k),k)$, we obtain a decreasing sequence of algebraic sets defined over k,

$$V \supseteq A^1 \supseteq A^2 \supseteq \cdots.$$

Recall that k is assumed perfect.

Proposition 1. *If $n = \dim(V)$, then $A^{n+1}(V,k) = A^{n+r}(V,k)$ for all $r \geqq 1$. Further $A^{n+1}(V,k) = A(V,k)$.*

Proof. Suppose $A^s(V,k) = A^s \neq A^{s+1}$. Then there is some k-component W^s of A^s which is not absolutely irreducible. $W^s \subseteq A^{s-1}$, but W^s is not a k-component of A^{s-1}, by the construction of A^s from A^{s-1}. Hence W^s is a proper subset of some k-component W^{s-1} of A^{s-1}, and W^{s-1} is not absolutely irreducible. In this manner we obtain a strictly increasing sequence of k-varieties, and since $0 \leqq \dim(W^s) \leqq \dim(V) - s$, $s \leqq \dim(V)$. We now show that $A^{n+1} = A$. Since $A^{n+2} = A^{n+1}$, every k-component of A^{n+1} is absolutely irreducible, and $A^{n+1} \subseteq A$. If now U is a subvariety of V, absolutely irreducible, and defined over k, then $U \subseteq V_{i'j'}$ for some i', j'. Since U is defined over k, $U \subseteq V_{i'j}$ for all j, and hence $U \subseteq \cap V_{i'j} \subseteq A^1(V,k)$. By induction, $U \subseteq A^{n+1}(V,k)$.

Corollary. *$A(V,k)$ is an algebraic set, being the union of a finite number of varieties defined over k.*

2. Change of ground field. If L is any field containing k, V is defined over L, and $A(V,L) \supseteq A(V,k)$. We shall show below that $A(V,L)$ depends only on $L \cap K$, where K is some finite extension of k, depending on V. If $V = \cup V_i = \cup V_{ij}$ is the representation of V as the union of its k-components and components, let K_i be the normal extension of k over which all of the V_{ij} are defined, and let $\mathbf{K}^1(V,k)$ be the composite of the K_i.

Proposition 2. *If L is any field containing k, then*

$$A^1(V,L) = A^1(V, L \cap \mathbf{K}^1(V,k)).$$

Proof. We first note the following two facts:

(I) If f is a polynomial in one variable over k, K the splitting field of f over k, then the irreducible factors of f in any extension field L of k coincide with those in $L \cap K$. Hence if α is a root of f,

$$[L(\alpha):L] = [(L \cap K)(\alpha): L \cap K].$$

(II) If V is a k-variety, and K is the normal field of definition over k of the absolutely irreducible components of V, then for any L containing k, the L-components of V coincide with the $(L \cap K)$-components.

(I) is totally elementary, and (II) is deduced from (I) as follows. We need only show that the $(L \cap K)$-components of V remain irreducible over L. If U is an $(L \cap K)$-component of V, and $U_1, \cdots, U_r$ are the components of U, then U_1 has minimum field of definition K_1 over $L \cap K$, of degree r. The minimum field of definition of U_1 over L is K_1L and by (I) $[K_1L: L] = r$. Hence U_1 has r conjugates over L and U is irreducible over L. Proposition 2 is now an easy corollary of (II).

Having defined $\mathbf{K}^1(V,k)$, we define $\mathbf{K}^t(V,k)$ by induction. Let $k_1, \cdots, k_m$ be all intermediate fields between k and $\mathbf{K}^{t-1}(V,k)$. Then $\mathbf{K}^t$ is defined to be the composite of the fields $\mathbf{K}^1(\mathbf{A}^{t-1}(V,k_j),k_j)$, with $j = 1, \cdots, m$. By a simple induction, we obtain from Proposition 2:

$$\mathbf{A}^t(V,L) = \mathbf{A}^t(V, L \cap \mathbf{K}^t(V,k)).$$

Setting $\mathbf{K}(V,k) = \mathbf{K}^{n+1}(V,k)$, we obtain:

PROPOSITION 3. *For any algebraic set V defined over k, there is a finite extension $\mathbf{K}(V,k)$ of k, such that for any field L containing k, $\mathbf{A}(V,L) = \mathbf{A}(V, L \cap \mathbf{K}(V,k))$.*

3. Reduction (mod $\mathfrak{p}$). We now assume that the field k is an algebraic number field. $k_{\mathfrak{p}}$, $\mathfrak{o}_{\mathfrak{p}}$, and $\bar{k}_{\mathfrak{p}}$ denote a p-adic completion of k, the ring of integers of $k_{\mathfrak{p}}$, and the residue class field of $k_{\mathfrak{p}}$. If V is an algebraic set defined over k, then, for each $\mathfrak{p}$, we may define the reduced algebraic set (mod $\mathfrak{p}$), which we shall denote by $\mathfrak{p}(V)$. For the details of this construction, see [8], and Chapter III of [9]. $\mathfrak{p}(V)$ does not depend essentially on the field $k_{\mathfrak{p}}$, being the same, for instance, if the reduction is carried out over a finite extension of $k_{\mathfrak{p}}$. We now collect in one proposition some relevant properties of reduction (mod $\mathfrak{p}$). By "for almost all $\mathfrak{p}$" we shall mean for all but a finite number of $\mathfrak{p}$.

PROPOSITION 4. *If U and V are algebraic sets over $k_{\mathfrak{p}}$, then:*

(i) $\mathfrak{p}(U \cup V) = \mathfrak{p}(U) \cup \mathfrak{p}(V)$.

(ii) *If all components of V are of dimension r, then either $\mathfrak{p}(V)$ is empty, or has all components of dimension r.*

If now U and V are defined over k,

(iii) $\mathfrak{p}(U \cap V) = \mathfrak{p}(U) \cap \mathfrak{p}(V)$ *for almost all* $\mathfrak{p}$.

(iv) $\mathfrak{p}(V)$ *is non-empty for almost all* $\mathfrak{p}$.

(v) *If V is a variety, then $\mathfrak{p}(V)$ is also a variety for almost all $\mathfrak{p}$.*

The proofs of these results are all found in [8] or [9]. (i) and (ii) are Propositions 18 and 19 of [8], (iii) and (iv) are Propositions 19 and 17 of [9, Ch. III], and (v), originally proved by E. Noether in [6], is Theorem 26 of [8].

PROPOSITION 5. *Let V be a k-variety, such that V is irreducible aver $k_{\mathfrak{p}}$ for all primes in some set $\mathcal{S}$. Then for almost all $\mathfrak{p}$ in $\mathcal{S}$. $\mathfrak{p}(V)$ is irreducible over $\bar{k}_p$. If V has m components, then so will $\mathfrak{p}(V)$, for almost all p.*

Proof. If V has components $V_1, \cdots, V_m$, it follows from Proposition 4 that, for almost all $\mathfrak{p}$, $\mathfrak{p}(V_1), \cdots, \mathfrak{p}(V_m)$ will be distinct and absolutely irreducible. For $\mathfrak{p}$ in $\mathcal{S}$, the isomorphism over $k_{\mathfrak{p}}$ which takes V_i to V_j induces an isomorphism over $\bar{k}_{\mathfrak{p}}$ taking $\mathfrak{p}(V_i)$ to $\mathfrak{p}(V_j)$. Hence for almost all $\mathfrak{p}$ in $\mathcal{S}$, $\mathfrak{p}(V)$ is the union of conjugate varieties over $\bar{k}_{\mathfrak{p}}$.

THEOREM 1. *Let V be an algebraic set defined over the algebraic number field k. For almost all $\mathfrak{p}$-adic completions $k_{\mathfrak{p}}$ of k, $\mathfrak{p}(\boldsymbol{A}(V, k_{\mathfrak{p}})) = \boldsymbol{A}(\mathfrak{p}(V), \bar{k}_{\mathfrak{p}})$.*

Proof. We need only show that $\mathfrak{p}(\boldsymbol{A}^1(V, k_{\mathfrak{p}})) = \boldsymbol{A}^1(\mathfrak{p}(V), \bar{k}_{\mathfrak{p}})$. In the formation of $\boldsymbol{A}^1(V, k_{\mathfrak{p}})$ a finite sequence of operations occurs, which are: decompositions into $k_{\mathfrak{p}}$-components, decompositions into absolute components, formation of intersections and of unions of algebraic sets. Since this sequence of operations depends not on $k_{\mathfrak{p}}$ itself, but only on $k_{\mathfrak{p}} \cap \boldsymbol{K}^1(V, k)$, only a finite collection of operations need be carried out in the construction of $\boldsymbol{A}^1(V, k_{\mathfrak{p}})$ for all $\mathfrak{p}$ taken together. Here it clearly makes no difference whether we regard $k_{\mathfrak{p}}$ as a $\mathfrak{p}$-adic completion of k, or of $k_{\mathfrak{p}} \cap \boldsymbol{K}^1(V, k)$. Propositions 4 and 5 now guarantee that each of the operations in the construction of $\boldsymbol{A}^1(V, k_{\mathfrak{p}})$ commutes with reduction $(\bmod\, p)$ for almost all p.

4. Applications. Theorem 1 allows us to use residue class field information in the following way.

THEOREM 2. *Let V be an algebraic set defined over the number field k, such that for almost all $\mathfrak{p}$, $\mathfrak{p}(V)$ contains a subvariety defined over $\bar{k}_{\mathfrak{p}}$ (i.e. $\boldsymbol{A}(\mathfrak{p}(V), \bar{k}_{\mathfrak{p}})$ is non-empty). Then there is a finite collection of subvarieties U_j of V, such that for almost all $\mathfrak{p}$, some U_j is defined over $k_{\mathfrak{p}}$.*

Proof. By Theorem 1, $\boldsymbol{A}(V, k_{\mathfrak{p}})$ will be non-empty for almost all $\mathfrak{p}$, and by Proposition 3, $\boldsymbol{A}(V, k_{\mathfrak{p}}) = \boldsymbol{A}(V, k_{\mathfrak{p}} \cap \boldsymbol{K}(V, k))$. Since $\boldsymbol{K}(V, k)$ is a finite extension of k, the result follows.

But, as noted by Lang in [4], a variety V defined over the number field k will necessarily have rational points over $k_{\mathfrak{p}}$ for almost all $\mathfrak{p}$. To see this, we first take hyperplane sections of V to reduce to the case of a curve C defined over k. By choosing the hyperplane sections to be sufficiently general, we may insure that the curve C is absolutely irreducible ([3], p. 212). For almost all $\mathfrak{p}$, $\mathfrak{p}(C)$ will be an absolutely irreducible curve of some fixed degree d, which is defined over $\bar{k}_{\mathfrak{p}}$. The number of singular points on $\mathfrak{p}(C)$ is the same for almost all $\mathfrak{p}$, while the Riemann hypothesis for function fields implies that the number of points of $\mathfrak{p}(C)$ which are rational over $\bar{k}_{\mathfrak{p}}$ increases at least as some multiple of $(q)^{\frac{1}{2}}$, where q is the number of elements of $\bar{k}_{\mathfrak{p}}$ (see [5]). Hence we are assured of non-singular points on $\mathfrak{p}(C)$, rational over $\bar{k}_{\mathfrak{p}}$, for almost all $\mathfrak{p}$. But by Hensel's Lemma a non-singular point of $\mathfrak{p}(C)$ with coordinates in $\bar{k}_{\mathfrak{p}}$ comes from a point of C with coordinates in $k_{\mathfrak{p}}$ (see [7]).

THEOREM 3. *Let the hypotheses of Theorem 2 hold. Then V contains a rational point in $k_{\mathfrak{p}}$ for almost all $\mathfrak{p}$.*

Proof. By Theorem 2, for almost all $\mathfrak{p}$, there is a subvariety U_j of V which is defined over $k_{\mathfrak{p}} \cap \boldsymbol{K}(V, k)$. Hence by the above remarks U_j will have a rational point in almost all $\mathfrak{p}$-adic completions of $k_{\mathfrak{p}} \cap \boldsymbol{K}(V, k)$ and $k_{\mathfrak{p}}$ is isomorphic to a completion of $k_{\mathfrak{p}} \cap \boldsymbol{K}(V, k)$.

By considering a projective algebraic set in P^n as an affine set in S^{n+1}, and reading "not equal to (0)" for "non-empty" we may prove Theorems 1, 2, and 3 for projective varieties. In particular we have

THEOREM 3′. *If V is a projective algebraic set, defined over the number field k, such that $\mathfrak{p}(V)$ contains a subvariety defined over $\bar{k}_{\mathfrak{p}}$ for almost all $\mathfrak{p}$, then V has rational points in $k_{\mathfrak{p}}$ for almost all $\mathfrak{p}$.*

If $\mathfrak{p}(V)$ has a rational point over $\bar{k}_{\mathfrak{p}}$, then $\boldsymbol{A}(\mathfrak{p}(V), \bar{k}_{\mathfrak{p}})$ is non-empty, and we may apply Theorem 3 or 3′. That this is often the case is guaranteed by the theorem of Chevalley [2].

PROPOSITION 6. *Let $\bar{f}_1, \cdots, \bar{f}_t$ be polynomials in n variables over the finite field $\bar{k}$, of degrees $d_1, \cdots, d_t$. If $n > \sum d_j$ and all $\bar{f}_j$ have constant term equal to zero, then the $\bar{f}_j$ have a common zero, other than the trivial one, in $\bar{k}$.*

THEOREM 4. *Let $f_1, \cdots, f_t$ be polynomials in n variables over the number field k, of degrees $d_1, \cdots, d_t$. If $n > \sum d_j$ and all f_j have zero constant term, then the f_j have a non-trivial common zero in $k_{\mathfrak{p}}$ for almost all $\mathfrak{p}$.*

Proof. Let V be the algebraic set over k which corresponds to the ideal generated by the f_j. Then the assertion of the theorem is equivalent to the statement that V has rational points in $k_{\mathfrak{p}}$ for almost all $\mathfrak{p}$. By Proposition 4(iii), for almost all $\mathfrak{p}$, $\mathfrak{p}(V)$ is the algebraic set corresponding to the ideal generated by the $\bar{f}_j$, where the coefficients of $\bar{f}_j$ are obtained from those of f_j by reduction $(\operatorname{mod} p)$. Proposition 6 then implies that $A(\mathfrak{p}(V), \bar{k}_{\mathfrak{p}})$ will contain points other than the origin (0), for almost all $\mathfrak{p}$. We then proceed as in the proofs of Theorems 3 and 3'.

COROLLARY. *If f is a homogeneous form of degree d in $d+1$ variables over the number field k, then f has a non-trivial zero in $k_{\mathfrak{p}}$ for almost all $\mathfrak{p}$.*

5. Remarks. It can easily be seen that the requirement in the Corollary to Theorem 4 that the form f have at least $d+1$ variables is necessary. For let L be a cyclic extension of the number field k, of degree d. Then, for all $\mathfrak{p}$ in some infinite set $\mathcal{S}$, $L_{\mathfrak{P}}$ is of degree d over $k_{\mathfrak{p}}$. Let $\{b_1, \cdots, b_d\}$ be a basis for L over k, and hence also a basis for $L_{\mathfrak{P}}$ over $k_{\mathfrak{p}}$ for all $\mathfrak{p}$ in $\mathcal{S}$. Set $g(x) = \sum b_i x_i$, and let f be equal to the product of the d conjugates of g over k. Then f is a homogeneous form of degree d in d variables with coefficients in k. If now $a_1, \cdots, a_d$ are any elements of the field $k_{\mathfrak{p}}$, $\mathfrak{p} \in S$, then $f(a) = N(\sum a_i b_i)$, where N is the norm from $K_{\mathfrak{P}}$ to $k_{\mathfrak{p}}$. Since the norm of a non-zero element is never zero, f has no zero in $k_{\mathfrak{p}}$, except for the trivial one.

It seems certain that Theorem 4 also holds when k is a field of algebraic functions in one variable over a finite constant field. But we have been unable to prove this in general, and indeed the method of proof given here breaks down in this case. Proposition 1 is of course false, since now k has inseparable extensions, and we give the following simple counter example to Theorem 3'. Let $k = \mathbf{Z}_2(\xi)$, the rational function field over a field with two elements. Let V be defined by the form $x_1^2 + \xi x_2^2 = 0$. If $L = k(\sqrt{\xi})$, then

$L_{\mathfrak{P}}$ is a totally ramified extension of $k_{\mathfrak{p}}$ of degree two for all $\mathfrak{p}$. Hence V has no rational points in $k_{\mathfrak{p}}$ for *all* $\mathfrak{p}$. But $\mathfrak{p}(V)$ *has* rational points in $\bar{k}_{\mathfrak{p}}$ for all $\mathfrak{p}$.

Ax and Kochen have recently obtained a totally different proof of Theorem 4.

HARVARD UNIVERSITY.

REFERENCES.

[1] B. J. Birch and D. J. Lewis, "p-adic forms," *Journal of the Indian Mathematical Society*, vol. 23 (1959), pp. 11-32.

[2] C. Chevalley, "Demonstration d'une hypothese de M. Artin," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 11 (1935), pp. 73-75.

[3] S. Lang, *Introduction to Algebraic Geometry*, Interscience Tracts No. 5, Interscience, New York, 1958.

[4] ———, "Some theorems and conjectures in diophantine equations," *Bulletin of the American Mathematical Society*, vol. 66 (1960), pp. 240-249.

[5] S. Lang and A. Weil, "Number of points of varieties in finite fields," *American Journal of Mathematics*, vol. 76 (1954), pp. 819-827.

[6] E. Noether, "Eliminationstheorie und allgemeine Idealtheorie," *Mathematische Annalen*, vol. 90 (1923), pp. 229-261.

[7] P. Samuel, "Remarques sur le lemme de Hensel," *Proceedings, International Congress of Mathematicians*, 1954, Amsterdam, vol. 2, pp. 63-64.

[8] G. Shimura, "Reduction of algebraic varieties with respect to a discrete valuation of the basic field," *American Journal of Mathematics*, vol. 77 (1955), pp. 134-176.

[9] G. Shimura and Y. Taniyama, *Complex Multiplication of Abelian Varieties*, Publications of the Mathematical Society of Japan, Tokyo, 1961.

INTEGRAL REPRESENTATIONS OF QUADRATIC FORMS IN CHARACTERISTIC 2.

By C. R. Riehm.*

A quadratic form q over an integral domain R is said to be represented by another such form Q if there is a linear substitution for the variables of Q which yields q; symbolically $q \to Q$. If the substitution is invertible (in R), q is equivalent to Q, $q \cong Q$. A foundation for these classical problems when R is a field of characteristic 2 was given by Arf in [1]; the theory he propounded was similar to that given earlier by Witt [9] for characteristic $\neq 2$. In particular, Arf gave a complete set of invariants for forms over local fields F of characteristic 2, i. e. power series fields in one variable over a finite field with 2^n elements. In [2], Arf gave a partial solution to the local integral equivalence problem, that is to say when R is the ring $\mathfrak{o} \subset F$ of integral power series over the finite field: $R = \{\sum_{i \geqq 0} a_i x^i\}$. Sah [8] completed this theory with a solution similar to O'Meara's [4] in the characteristic 0 case.

The local field representation problem can be deduced from Arf's results; this has been done by Sah [7]. In this paper we give a partial solution to the local integral representation problem. The techniques and results are similar to those for the characteristic 0 case [6], although they are somewhat more general, and the conditions are more satisfactory. The results are expressed in the language of "lattices" rather than that of quadratic forms; the form is then a certain mapping of a free $\mathfrak{o}$-module $L(Q)$ into $\mathfrak{o}$ (or possibly F).

With each quadratic form Q, there is associated a bilinear form $B(x,y) = Q(x+y) + Q(x) + Q(y)$. The scale $\mathfrak{s}Q$ is the ideal $\{B(x,y) \mid x, y \text{ in } L(Q)\}$. The norm group $\mathfrak{g}Q$ is the additive subgroup of F generated by $\{Q(x) \mid x \in L(Q)\}$. With each non-defective form Q, we associate a "dual" form $Q^{\#}$ (cf. 1.29). We may assume that $\nu =$ (the number of variables of Q) minus (the number of variables of q) is $\geqq 0$. The basic result for our theory of integral representation is Proposition 4.2 which states that, *in certain*

Received February 20, 1964.

* The research for this paper was partially supported by the National Science Foundation under grant GP-1656.

circumstances when $\nu=0$, *necessary and sufficient conditions for* $q\to Q$ *are simply field equivalence,* $\mathfrak{s}q\subseteq\mathfrak{s}Q$, $\mathfrak{g}q\subseteq\mathfrak{g}Q$, *and* $\mathfrak{g}Q^{\#}\subseteq\mathfrak{g}q^{\#}$. The most general results given here for $\nu=0,1,2$ and $\nu\geqq 3$ are Theorems 4.8, 4.11, 4.10, and 4.9 respectively. The last theorem, for example, says that *when "q is well below Q"* (cf. 4.3) *and when* $\nu\geqq 3$, *the conditions* $\mathfrak{s}q\subseteq\mathfrak{s}Q$ *and* $\mathfrak{g}q\subseteq\mathfrak{g}Q$ *are necessary and sufficient for* $q\to Q$. These four theorems solve *completely* the representation problem $q\to Q$ when Q is "modular" (cf. 1.14).

In order to facilitate the development of the integral representation theory, a theory of maximal lattices analogous to that of Eichler [3] is developed in § 2. It culminates in Theorem 2.4 which states that two $\mathfrak{a}$-maximal lattices on the same space are equivalent. There are also useful results concerning the existence of maximal lattices containing (Lemma 2.7) or contained in (Lemma 2.11) a given lattice.

Finally Theorem 1.12 gives the structure of the group of non-defective spaces over a local field of characteristic 2. This group was introduced by Arf [1] in analogy to the well-known Witt group for characteristic $\neq 2$.

§ 1. Preliminaries.

1.1. The field F and its ring of integers $\mathfrak{o}$. The field F is the power series field $f\langle\langle\pi\rangle\rangle$ in one variable π over a finite field f of characteristic 2. The rational integer $\operatorname{ord}\alpha$ is the *order* of $\alpha\in F$, ($\operatorname{ord}0=\infty$) and $|\alpha|=c^{\operatorname{ord}\alpha}$ for some (fixed) real number c with $0<c<1$, is the corresponding multiplicative *valuation.* The ring of integers is $\mathfrak{o}=f[[\pi]]=\{\alpha\in F\mid\operatorname{ord}\alpha\geqq 0\}$ and its unique maximal ideal is $\mathfrak{p}=\{\alpha\mid\operatorname{ord}\alpha>0\}$. The group of units is denoted by $\mathfrak{u}$. If α and β are in F and $\operatorname{ord}\alpha\equiv\operatorname{ord}\beta\pmod 2$, we write $\alpha\sim\beta$, otherwise $\alpha\not\sim\beta$. In particular $\alpha\not\sim 0$ if $\alpha\neq 0$. If $\mathfrak{a}=a\mathfrak{o}$ and $\mathfrak{b}=b\mathfrak{o}$ are ideals in F, then $\mathfrak{a}\sim\mathfrak{b}$, $(\mathfrak{a}\sim\alpha)$, means $a\sim b$, $(a\sim\alpha)$.

Let F^2 (resp. $\mathfrak{o}^2$) be the set of squares in F (resp. $\mathfrak{o}$). The *quadratic defect* $\mathfrak{d}(\alpha)$ of $\alpha\in F$ is the ideal $\gamma\mathfrak{o}$ where $\alpha=\beta^2+\gamma$, $\gamma\not\sim 1$. It is the smallest ideal $\mathfrak{a}$ such that $\alpha\in F^2+\mathfrak{a}$, and has the following properties: $\mathfrak{d}(\alpha)\subseteq\alpha\mathfrak{o}$, $\mathfrak{d}(\alpha)=\alpha\mathfrak{o}\Longleftrightarrow\operatorname{ord}\alpha$ is odd, $\mathfrak{d}(\delta^2\alpha)=\delta^2\mathfrak{d}(\alpha)$, $\mathfrak{d}(\xi^2+\eta)=\eta\mathfrak{o}$ if $\operatorname{ord}\eta$ is odd.

A *norm group* is a finitely generated $\mathfrak{o}^2$-module in F. Let $\mathfrak{g}$ be a non-zero norm group and let $a\in\mathfrak{g}$ be of maximal value $|a|$. If $c\in\mathfrak{g}$, write $ca^{-1}=\xi^2+\eta$ where $\eta\not\sim 1$; then $a\eta=c-\xi^2a\in\mathfrak{g}$. Thus it is clear that if we choose $b\in\mathfrak{g}$ such that $b\not\sim a$ and $|b|$ is maximal for all such elements, then $\mathfrak{g}=a\mathfrak{o}^2+b\mathfrak{o}^2=a\mathfrak{o}^2+b\mathfrak{o}$ (b may be zero). Elements a and b with these properties are called *norm* and *base generators* of $\mathfrak{g}$ respectively. Whenever

a norm group appears in the form $a\mathfrak{o}^2 + b\mathfrak{o}$, it will be tacitly assumed that a is a norm generator and that b is a base generator.

It follows directly from the definition that the sum $\mathfrak{g}_1 + \mathfrak{g}_2$ and the intersection $\mathfrak{g}_1 \cap \mathfrak{g}_2$ of two norm groups are themselves norm groups. Let us show how to determine generators for them; suppose $\mathfrak{g}_i = a_i\mathfrak{o}^2 + b_i\mathfrak{o}$. If $0 \neq |a_1| \geqq |a_2|$, then $a = a_1$ is clearly a norm generator of $\mathfrak{g}_1 + \mathfrak{g}_2$, and it can be shown that a base generator b is determined by

$$b\mathfrak{o} = a^{-1}(\mathfrak{d}(ab_1) + \mathfrak{d}(aa_2) + \mathfrak{d}(ab_2)).$$

A norm generator of $\mathfrak{g}_1 \cap \mathfrak{g}_2$ is an element of largest value therein. Now it is easy to see that the largest ideal contained in a norm group with base generator b' is $b'\mathfrak{p}^{-1}$. From this it follows that a base generator b for $\mathfrak{g}_1 \cap \mathfrak{g}_2$ is given by $b\mathfrak{o} = b_1\mathfrak{o} \cap b_2\mathfrak{o}$.

The mapping $\wp : F \rightarrow F$ given by $\wp(\alpha) = \alpha^2 + \alpha$ is a homomorphism of additive groups with kernel $\{0, 1\}$. Thus $[f : \wp(f)] = 2$ and so we can choose representatives 0 and ρ of $f \bmod \wp(f)$; the quantity ρ will be fixed from now on. Now let $\mathfrak{D}$ be the additive group

$$\{\pi^{-1}b^2 + \omega \mid b \in f[\pi^{-1}], \omega = 0 \text{ or } \rho\}$$

where $f[\pi^{-1}]$ is the ring of polynomials in π^{-1} over f. Some elementary calculations and Hensel's Lemma (cf. Lemma 1.1 in [8]) show that F is the direct sum of $\mathfrak{D}$ and $\wp(F)$:

$$F = \mathfrak{D} \oplus \wp(F).$$

Note that if $\delta \in \mathfrak{D}$, then $\operatorname{ord} \delta$ is an odd negative integer, zero, or ∞ if $\delta = 0$. Also

$$(1) \qquad |\delta| = \inf_{\alpha \in F} \{|\delta + \wp(\alpha)|\}.$$

1.2. Quadratic spaces and lattices. A *quadratic space* V is a finite-dimensional vector space over F with a map $Q: V \rightarrow F$ such that $Q(\alpha x) = \alpha^2 Q(x)$, and $B(x, y) = Q(x + y) + Q(x) + Q(y)$ is bilinear, where $x, y \in V$ and $\alpha \in F$. A *lattice* L is a finitely generated $\mathfrak{o}$-module in such a space. The space spanned by L is written FL, and we say that L is *on* V if $FL = V$. If $e_1, \cdots, e_n$ is any basis of V and $x = \sum_i \xi_i e_i$, then

$$Q(x) = \sum_i \xi_i^2 Q(e_i) + \sum_{i<j} \xi_i \xi_j B(e_i, e_j)$$

is a quadratic form in $\xi_1, \cdots, \xi_n$. Any lattice L has a basis

$$L = \mathfrak{o}e_1' + \cdots + \mathfrak{o}e_m' \text{ (direct)},$$

and $FL = Fe_1' + \cdots + Fe_m'$. We put $\dim L = m$ ($= \dim FL$).

Orthogonality in V (resp. L) is defined by the bilinear form B. If $X \subset V$ (resp. L) then X^* is the subspace of V (resp. sublattice of L) orthogonal to X. If L is the direct sum of the mutually orthogonal sublattices J and K, then we write $L = J \perp K$; similarly for spaces. Note that $F(J^*) = (FJ)^*$. If $\dim L^* > 0$, L is called *defective*; otherwise L is *non-defective*. Similarly for spaces. For the problem of representation in which we are interested, it suffices to restrict the investigations to lattices L with the property that

$$\{x \in L^* \mid Q(x) = 0\} = 0$$

and we shall accordingly assume in this paper that all lattices (and spaces) given are non-degenerate in this sense. With this assumption, $\dim L^* \leqq [\mathfrak{o}:\mathfrak{o}^2] = 2$, (index of $\mathfrak{o}^2$-modules).

If $L = \mathfrak{o}x$ with $Q(x) = a$, then we write $L \cong \langle a \rangle$, and also $FL \cong \langle a \rangle$—this similarity of notation does not cause confusion since it is always clear from the context whether $\langle a \rangle$ represents a lattice or a space. If L is any totally defective lattice ($L^* = L$) of dimension one, then $L \cong \langle a \rangle$ for some $a \neq 0$. If L is totally defective of dimension 2, then $L \cong \langle a \rangle \perp \langle b \rangle$ for suitable non-zero a and b; in fact, $Q(L)$ is a norm group, and a and b can be chosen to be *any* norm and base generators respectively. The same expressions hold for spaces. If $L = \mathfrak{o}x + \mathfrak{o}y$ with $Q(x) = a$, $B(x, y) = b \neq 0$, and $Q(y) = c$, then we write $L \cong (a^b c)$; if $b = \pi^i$, write $L \cong \mathfrak{A}_i(a, c)$, and finally if $i = 0$, we often write $L \cong \mathfrak{A}(a, c)$. In [8], it is shown that any non-defective lattice L is an orthogonal sum of these binary lattices: $L \cong \mathfrak{A}_{i(1)}(a_1, c_1) \perp \cdots \perp \mathfrak{A}_{i(n)}(a_n, c_n)$. This will be called a *quasi-diagonal splitting*. Since any non-defective space V supports such a lattice L, we have *a posteriori* $V \cong F\mathfrak{A}_{i(1)}(a_1, c_1) \perp \cdots \perp F\mathfrak{A}_{i(n)}(a_n, c_n)$. Note that if we take any direct splitting $L = L^\ddagger \oplus L^*$, then it is also an orthogonal splitting, $L = L^\ddagger \perp L^*$, and $L^\ddagger$ is non-defective. A lattice $L \cong \mathfrak{A}_i(0, 0)$ (or the space spanned by it) is called a hyperbolic plane and is usually denoted by $\mathfrak{H}_i$ ($F\mathfrak{H}_i$ for the space). An orthogonal sum of hyperbolic planes, or 0, is called a *hyperbolic* lattice, and its space a hyperbolic space. A vector x with $Q(x) = 0$ is *isotropic*, and any lattice (space) containing such vectors is called *isotropic*. An isotropic space V always splits off a hyperbolic plane [1]: $V \cong F\mathfrak{H} \perp U$. If L is non-degenerate, then L^* is not isotropic because of the non-degeneracy assumption; similarly for spaces.

1.3. Arf invariant. Let L be a non-defective lattice; by 1.2, $L \cong \perp_{1 \leqq i \leqq n} (a_i^{b_i} c_i)$. The coset $\sum_i a_i c_i b_i^{-2} + \wp(F)$ is an invariant of FL, the *Arf invariant*, [1]. By 1.1, there is a uniquely determined representative

$\Delta(L) \in \mathfrak{D}$ of this coset. We shall generally prefer the invariant thus provided to the Arf invariant itself, i.e. to the coset $\Delta(L) + \wp(F)$. If J and K are non-defective lattices, then clearly $\Delta(J \perp K) = \Delta(J) + \Delta(K)$. A splitting of a non-defective space V into binary spaces leads to the Arf invariant of V in the same way, and thence to the invariant $\Delta(V)$. Trivially $\Delta(L) = \Delta(FL)$. If $V \cong F\mathfrak{H}$, then clearly $\Delta(V) = 0$; conversely [1], if V is binary and $\Delta(V) = 0$, then $V \cong F\mathfrak{H}$.

1.4. Representation and isometry of spaces. A *representation* is a linear map $\sigma\colon V \to W$ between two quadratic spaces which preserves the quadratic form: $Q(\sigma x) = Q(x)$. It follows that B is also preserved. Since V is non-degenerate, σ is injective. Also it is clear that $\sigma\colon V^* \to W^*$. To say that V is represented by W, $V \to W$, is the same as saying that W contains a copy of V as a subspace.

If the representation $V \to W$ is an isomorphism, we shall write $V \cong W$; V is *isometric* to W. The Arf invariant, the dimension, and the Clifford algebra were first used by Arf [1] to characterize isometric spaces. We shall not use this theory explicitly here. In the same paper, the analogy to Witt's Theorem of the characteristic $\neq 2$ theory, is proved: *If $V_1 \perp V_2 \cong V_1' \perp V_2'$ and $V_1 \cong V_1'$ is non-defective, then $V_2 \cong V_2'$.*

Another basic result is Zusatz 1, p. 153 in [1]: *If V is a non-defective binary space and $a \in Q(V)$, $a \neq 0$, and $\delta \equiv \Delta(V) \bmod \wp(F)$, then $V \cong F\mathfrak{A}(a, \delta a^{-1})$.*

The phenomenon $|\alpha + \beta| = |\alpha|$ if $|\alpha| > |\beta|$ is referred to as *domination*. An important application of it is the following. Suppose that $V \cong F\mathfrak{A}(a, 0)$ in the basis $V = Fx + Fy$, and suppose that $\xi x + \eta y \in V$, $|Q(\xi x)| > |B(\xi x, \eta y)|$. Then $|Q(\xi x + \eta y)| = |Q(\xi x)|$. This situation arises, for example, when $a \notin \mathfrak{o}$, ξ and $\eta \in \mathfrak{o}$, and $|\xi| > |a^{-1}|$. The next lemma provides a similar (and stronger) conclusion for the case $\dim V = 2$, $\Delta(V) \neq 0$. *The word* domination *will be used throughout the sequel to describe all of these situations.*

LEMMA 1.5. *Let V be a non-defective binary space. Let $x, y \in V$ with $|Q(x)Q(y)B(x,y)^{-2}| = |\Delta(V)| \neq 0$. Then*

$$|Q(x+y)| = |Q(x) + Q(y) + B(x,y)| = \sup\{|Q(x)|, |Q(y)|\}.$$

Proof. It is easy to see that $|Q(x+y)| \leqq \sup\{|Q(x)|, |Q(y)|\}$ since

$\Delta(V) \neq 0$. Suppose $|Q(x+y)| < |Q(x)|$. Then $V = F(x+y) + Fy$ and since $B(x+y,y) = B(x,y)$,

$$Q(x+y)Q(y)B(x,y)^{-2} \equiv Q(x)Q(y)B(x,y)^{-2} \bmod \mathfrak{p}(F)$$

which contradicts $|\Delta(V)| = \inf_{\alpha \in F}\{|\Delta(V) + \mathfrak{p}(\alpha)|\}$.

This lemma will be used most often in the following form: if $V \cong F\mathfrak{A}(a, \delta a^{-1})$ with $0 \neq \delta \in \mathfrak{D}$, then

$$|\xi^2 a + \xi\eta + \eta^2 \delta a^{-1}| = \sup\{|\xi^2 a|, |\eta^2 \delta a^{-1}|\}.$$

LEMMA 1.6. $Q(F\mathfrak{A}(a, \rho a^{-1})) = a\mathfrak{u}F^2$.

Proof. First let us show that the equation

$$X^2 + XY + \rho Y^2 = \epsilon$$

has a solution $(X, Y) = (\xi, \eta)$ in F for any unit ϵ. Choose $\eta \in \mathfrak{u}$ such that $\rho\eta^2 \equiv \epsilon \bmod \mathfrak{p}$. The equation $X^2 + \eta X + (\rho\eta^2 + \epsilon) = 0$ has a solution $X = \xi$ by Hensel's Lemma. Thus $\xi^2 + \xi\eta + \rho\eta^2 = \epsilon$. It follows that $Q(F\mathfrak{A}(a, \rho a^{-1})) \supseteq a\mathfrak{u}F^2$, while the opposite inclusion is a consequence of domination.

LEMMA 1.7. *Let V be a quaternary non-isotropic space. Then*

$$V \cong F\mathfrak{A}(1, \rho) \perp F\mathfrak{A}(\pi, \rho\pi^{-1}).$$

Proof. We can write $V = U \perp W$ where $U \cong F\mathfrak{A}(a, \delta a^{-1})$ and $W \cong F\mathfrak{A}(b, \delta' b^{-1})$ with δ and δ' in $\mathfrak{D}$ (and non-zero). Suppose that $\delta \neq \rho$. Then $\delta' \neq \rho$ since otherwise W would represent an element which U does, by Lemma 1.6 and since $a \not\sim \delta a^{-1}$, and thus V would be isotropic. Therefore there is a vector x in W such that $Q(x) \sim \delta a^{-1}$. Choose $\alpha \in F$ such that $\alpha^2 Q(x) \equiv \delta a^{-1} \bmod \delta a^{-1}\mathfrak{p}$. If $U = Fy + Fz \cong F\mathfrak{A}(a, \delta a^{-1})$, the space $U' = F(\alpha x + z) + Fy$ splits V and clearly $|\Delta(U')| < |\Delta(U)|$. If $\Delta(U') \neq \rho$, we may repeat the process. Ultimately we get a binary space U_1 such that $V = U_1 \perp W_1$ and $\Delta(U_1) = \rho$. Since V is not isotropic, $\Delta(W_1) = \rho$ also. Therefore by Lemma 1.6, one of U_1 and V_1 is isometric to $F\mathfrak{A}(1, \rho)$ and the other to $F\mathfrak{A}(\pi, \rho\pi^{-1})$ as required.

COROLLARY 1.8. *A quaternary space V with $\Delta(V) \neq 0$ is isotropic.*

COROLLARY 1.9. *A quaternary space represents all of F.*

Proof. If it is isotropic, it splits off $F\mathfrak{H}$ which clearly represents F. If it is not isotropic, the corollary follows from the fact that $\langle 1 \rangle \perp \langle \pi \rangle$ represents F, and the lemma.

COROLLARY 1.10. *A space of dimension* $\geqq 5$ *is isotropic* (Satz 11, [1]).

Proof. Directly from Corollary 1.9.

Complete conditions for the representation of one space by another can be deduced from the results in [1], (see [7]). We shall give the conditions only for the cases which we shall need later.

THEOREM 1.11. *Let* V *be a non-defective space and* U *a non-zero space with* $\nu = \dim V - \dim U \geqq 0$. *Then* $U \to V$ *if and only if*

$$\nu = 0: \ U \cong V$$
$$\nu = 1: \ U^{\ddagger} \perp F\mathfrak{A}(a, \delta a^{-1}) \cong V$$
$$\textit{where } U \cong U^{\ddagger} \perp \langle a \rangle \textit{ and } \delta = \Delta(U^{\ddagger}) + \Delta(V).$$
$$\nu = 2: \ U^{*} = 0 \textit{ and } \Delta(U) = \Delta(V) \Rightarrow U \perp F\mathfrak{H} \cong V.$$
$$\nu \geqq 3: \textit{ no conditions.}$$

Proof. The proof of the necessity is straightforward and is omitted. The sufficiency is trivial for $\nu = 0$ and 1. Now suppose that $U = U^{*}$ is binary and V is quaternary. Split V into two non-defective binary spaces and choose x from one and y from the other with the property that neither $Q(x)$ nor $Q(y)$ is zero, and $Q(x) \notin Q(y) \cdot F^2$. Then

$$U \cong \langle Q(x) \rangle \perp \langle Q(y) \rangle \to V.$$

Now suppose that $\dim V = 6$. Then $V \cong W \perp F\mathfrak{H}$ by Corollary 1.10. Let $a \neq 0$, and $\delta \in \mathfrak{D}$. Then $F\mathfrak{H} \cong F\mathfrak{A}(a, 0)$, and since $Q(W) = F$ by Corollary 1.9, $V \cong W' \perp F\mathfrak{A}(a, \delta a^{-1})$. Thus it is clear that $U \to V$ if U is non-defective and $\nu \geqq 4$. Suppose $U \cong U^{\ddagger} \perp \langle a \rangle$, $U^{*} \cong \langle a \rangle$, and that $\nu \geqq 3$. Then $U^{\ddagger} \to V$ by the previous case, and so $V \cong U^{\ddagger} \perp U'$. But $Q(U') = F$ and so $\langle a \rangle \to U'$, whence $U \to V$. If $\nu \geqq 2$ and $\dim U^{*} = 2$, say $U = U^{\ddagger} \perp U^{*}$, then $U^{\ddagger} \to V$ by a case already treated, so that $V \cong U^{\ddagger} \perp W$. But by the first case dealt with, $U^{*} \to W$, and so $U \to V$ as required.

The only remaining case is $\nu = 2$, $U^{*} = 0$, and $\Delta(V) \neq \Delta(U)$. Then $U \to V \perp F\mathfrak{H}$, i.e. $V \perp F\mathfrak{H} \cong U \perp W$ for some non-defective quaternary space W. But $\Delta(W) = \Delta(U) + \Delta(V) + \Delta(F\mathfrak{H}) \neq 0$, and so $W \cong W' \perp F\mathfrak{H}$ by Corollary 1.8. By the Witt-Arf theorem, $V \cong U \perp W'$, whence $U \to V$ as required.

In [1], Arf defines an equivalence relationship $\approx$ among all quadratic spaces over F, as follows. Any space V has a splitting $V = H(V) \perp G(V)$ where $H(V)$ is a hyperbolic space and $G(V)$ is non-isotropic (or zero). We say that $V \approx W$ if $G(V) \cong G(W)$. The set of equivalence classes is a

monoid under the addition $\perp$. The set of equivalence classes consisting only of non-defective spaces is actually a group.

THEOREM 1.12. *The group $\mathfrak{G}$ of equivalence classes of non-defective spaces contains a subgroup $\mathfrak{H}$ isomorphic to the cyclic group Z_2 of two elements, and $\mathfrak{G}/\mathfrak{H}$ is the (weak) direct sum of a countable number of copies of Z_2.*

Proof. By Lemma 1.7 and Corollary 1.10, the only non-isotropic non-defective space of dimension > 2 is $V \cong F\mathfrak{A}(1,\rho) \perp F\mathfrak{A}(\pi,\rho\pi^{-1})$. It is easy to see that $V \perp V$ is a hyperbolic space, i.e. $V \perp V \approx 0$. Thus the subgroup $\mathfrak{H}$ consisting of the equivalence classes of V and 0 is isomorphic to Z_2. Also it is easy to check that $\Delta\colon \mathfrak{G} \to \mathfrak{D}$ is a group homomorphism, and that the kernel of Δ is $\mathfrak{H}$. Thus $\mathfrak{G}/\mathfrak{H} \cong \mathfrak{D}$. But $\mathfrak{D} = \pi^{-1}f[\pi^{-2}] + \{0,\rho\}$ which is isomorphic to a countable weak direct sum of Z_2.

LEMMA 1.13. *Let V be a non-defective space and suppose that x and y are two independent vectors in V such that $B(x,y) = 0$. Then there are vectors $x', y' \in V$ and a non-defective subspace V' (possibly zero) such that*

$$V = (Fx + Fx') \perp (Fy + Fy') \perp V'$$

with $B(x,x') = 1 = B(y,y')$.

Proof. Choose $x'' \in V$ such that $B(x,x'') = 1$ and split

$$V = (Fx + Fx'') \perp U.$$

Since $B(x,y) = 0$, and x and y are independent, $y = \alpha x + z$ with $0 \neq z \in U$. Choose $y' \in U$ with $B(z,y') = 1$, and let $W = Fy + Fy'$. Clearly $x \in W^*$; choose $x' \in W^*$ with $B(x,x') = 1$. This leads to the desired splitting, with V' the orthogonal complement of $Fx + Fx'$ in W^*.

1.14. Invariants of a lattice. The *scale* $\mathfrak{s}L$ of a lattice L is the ideal $B(L,L)$. The *norm* $\mathfrak{n}L$ is the ideal generated by $Q(L)$, $\mathfrak{n}L = Q(L)\cdot\mathfrak{o}$. It is easy to see that $\mathfrak{s}L \subseteq \mathfrak{n}L$. The *norm group* $\mathfrak{g}L$ of L is the smallest (additive) subgroup of F containing $Q(L)$; since $\mathfrak{o}^2\cdot Q(L) = Q(L)$, it is clearly a norm group in the sense of 1.1. Now if x and y are vectors in L with $B(x,y)\cdot\mathfrak{o} = \mathfrak{s}L$, and $\lambda \in \mathfrak{o}$, then

$$\lambda B(x,y) = Q(x+\lambda y) + Q(x) + \lambda^2 Q(y) \in \mathfrak{g}L$$

and therefore $\mathfrak{s}L \subseteq \mathfrak{g}L$. Since $Q(L) + \mathfrak{s}L$ is a group, we must have $\mathfrak{g}L = Q(L) + \mathfrak{s}L$. Thus $\mathfrak{s}L \subseteq \mathfrak{g}L \subseteq \mathfrak{n}L$. The expression $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$ can

be derived as follows (cf. 1.1). Let a be an element of maximal value $|a|$ in $Q(L)$. If $L = \mathfrak{o}x_1 + \cdots + \mathfrak{o}x_n$ and $\mathfrak{b} = \mathfrak{p}\mathfrak{s}L$ or $\mathfrak{s}L$ with $\mathfrak{b} \not\sim a$, then

$$\mathfrak{b}\mathfrak{o} = a^{-1} \sum_{i=1}^{n} \mathfrak{d}(Q(x_i)a) + \mathfrak{b}.$$

It is often convenient to normalize the quadratic form by defining a new quadratic form Q^α by $Q^\alpha(x) = \alpha Q(x)$, where $0 \neq \alpha \in F$. Then $B^\alpha = \alpha B$ also. This is called *scaling,* and is also done for spaces. If L is the lattice with the quadratic form Q, we write L^α for the same module with the form Q^α.

The operators $\mathfrak{s}$, $\mathfrak{n}$, and $\mathfrak{g}$ are all linear in the sense that, for example,

$$\mathfrak{g}(L_1 \perp L_2) = \mathfrak{g}L_1 + \mathfrak{g}L_2, \mathfrak{g}(L^\alpha) = \alpha\mathfrak{g}L.$$

A vector $x \in L$ is called *maximal* if $\pi^{-1}x \notin L$. By the "elementary divisor theorem" (81:11 in [5]), a maximal vector can be chosen as part of a basis for L, i.e. $L = \mathfrak{o}x + \mathfrak{o}y + \cdots + \mathfrak{o}z$ where the vectors $x, y, \cdots, z$ are independent.

A lattice L is $\mathfrak{a}$-*modular* ($\mathfrak{a}$ an ideal) if $B(x, L) = \mathfrak{a}$ for every maximal vector $x \in L$. Thus if L is $\mathfrak{p}^r$-modular, any quasi-diagonal splitting of L is of the form $L \cong \mathcal{A}_r(a_1, b_1) \perp \cdots \perp \mathcal{A}_r(a_n, b_n)$. And L 0-modular means that $L = L^*$, i.e. L is totally defective. An $\mathfrak{a}$-modular lattice L has scale $\mathfrak{s}L = \mathfrak{a}$.

Now let L be any lattice, and take a quasi-diagonal splitting for an orthogonal complement of L^*. By grouping these components according to their scales, we get a splitting

$$L = L_1 \perp L_2 \perp \cdots \perp L_T, \mathfrak{s}L_1 \supset \mathfrak{s}L_2 \supset \cdots \supset \mathfrak{s}L_T,$$

with each L_i modular. Such a splitting is called a *Jordan splitting.* If $L = K_1 \perp \cdots \perp K_S$ is any other such splitting, then $S = T$, and $\mathfrak{s}L_i = \mathfrak{s}K_i$, $\dim L_i = \dim K_i$ for $1 \leqq i \leqq T$; the proof of this invariance is entirely similar to the analogous situation in the characteristic $\neq 2$ theory (see 91:9 in [5]), and is omitted. There is an invariant derived from these which is useful. It is $\Sigma_i(L) = 0$ if $\mathfrak{s}L \subset \mathfrak{p}^i$, ($\mathfrak{p}^\infty = 0$), otherwise

$$\Sigma_i(L) = \dim L_1 + \cdots + \dim L_{n(i)}$$

where $n(i)$ is the largest integer ($\leqq T$) such that $\mathfrak{p}^i \subseteq \mathfrak{s}L_{n(i)}$. Thus, for example, $\Sigma_i(L) = 0$ for $i < \operatorname{ord}(\mathfrak{s}L)$, $\Sigma_i(L) = \dim L$ for $i \geqq \operatorname{ord}(\mathfrak{s}L_T)$.

1.15. Representation and isometry of lattices. We say that the lattice l is *represented* by L (in symbols $l \to L$) if there is a linear map $\sigma: l \to L$ which preserves the quadratic form (and therefore also the bilinear form). As in the case of spaces, the non-degeneracy assumption implies that σ is

injective. If σ is an isomorphism, we say that l is *isometric* to L (symbolically, $l \cong L$). Again, $l \to L$ if and only if L contains a replica of l.

The following theorem is of fundamental importance in this paper. It was proved by C.-H. Sah (Theorem 4.6 in [8]).

THEOREM. *Let L_1 and L_2 be modular lattices. Then $L_1 \cong L_2$ if and only if $FL_1 \cong FL_2$, $\mathfrak{s}L_1 = \mathfrak{s}L_2$, and $\mathfrak{g}L_1 = \mathfrak{g}L_2$.*

LEMMA 1.16. (i) *If L is a $\mathfrak{p}^i$-modular lattice ($i \neq \infty$) of dimension $\geqq 6$, then L splits off a hyperbolic plane: $L \cong K \perp \mathcal{H}_i$.*

(ii) *If H is a hyperbolic lattice and $H \perp L \cong H \perp K$, then $L \cong K$.*

(iii) *If P is a binary lattice with $\Delta(P) = 0$ and $\mathfrak{n}P = \mathfrak{s}P$, then $P \cong \mathcal{H}_i$.*

Proof. Lemma 4.4, Theorem 3.3, and Lemma 3.2, in [8].

LEMMA 1.17. *Let J be an $\mathfrak{a}$-modular sublattice of L, $\mathfrak{a} \neq 0$. Then J splits L, i.e. $L = J \perp K$, if and only if $B(J, L) \subseteq \mathfrak{a}$.*

Proof. See Lemma 1.3 in [8].

There are some obvious necessary conditions for $l \to L$. First of all, any such representation can be extended to a representation $Fl \to FL$ of spaces; thus we can assume, if we wish, that $Fl \subseteq FL$, or that $l \subset FL$. Next, the conditions $\mathfrak{s}l \subseteq \mathfrak{s}L$, $Q(l) \subseteq Q(L)$, and $\mathfrak{g}l \subseteq \mathfrak{g}L$ are necessary. We shall avoid $Q(l) \subseteq Q(L)$ since it is usually intractable, both theoretically and practically.

LEMMA 1.18. *Suppose $l \to L$. Then for all i, $\Sigma_i(L) \geqq \Sigma_i(l)$.*

Proof. We may suppose that $i \neq \infty$. Suppose $\phi: l \to L$, and split $L = J \perp K$ where $\mathfrak{s}K \subset \mathfrak{p}^i$ and $\Sigma_i(L) = \dim J$.

Define the linear map $\psi: l \to J$ by $\phi x = \psi x + x'$ where $x \in l$, $\psi x \in J$, and $x' \in K$. Clearly $B(x, y) = B(\phi x, \phi y) \equiv B(\psi x, \psi y) \bmod \mathfrak{p}^{i+1}$, which shows that $B(x, l) \supseteq \mathfrak{p}^i \Rightarrow \psi x \neq 0$. It follows readily that $\Sigma_i(l) \leqq \Sigma_i(L)$.

1.19. Canonical forms.

LEMMA 1.20. *Let L be a binary $\mathfrak{p}^i$-modular lattice ($i \neq \infty$), and let $a \in Q(L)$ be a norm generator of $\mathfrak{g}L$. Then*

$$L \cong \mathcal{A}_i(a, \Delta(L) a^{-1} \pi^{2i}).$$

Proof. See Corollary 4.2 in [8].

If $K \cong \mathcal{A}_i(a, b)$ with $\mathfrak{n}K = a\mathfrak{o}$ and $|ab\pi^{-2i}| = |\Delta(K)|$, then we shall write $K \cong \mathcal{B}_i(a, b)$; also $\mathcal{B}_0$ will usually be replaced by $\mathcal{B}$. Thus for example,

in 1.20 we have $L \cong \mathcal{B}_i(a, \Delta(L)a^{-1}\pi^{2i})$. Also if $\epsilon \in \mathfrak{u}$, $\mathcal{H}_i \cong \mathcal{B}_i(\epsilon\pi^i, 0)$ by 1.16 (iii).

LEMMA 1.21. *Let V be a binary space with $\Delta(V) \notin \mathfrak{o}$, and let $a \neq 0$ be any field element. Then there is a $\delta \in \Delta(V)^{-1} \cdot \mathfrak{o}$ such that $a(1+\delta) \in Q(V)$. And if $\delta \in \Delta(V)^{-1} \cdot \mathfrak{p}$, then $a \in Q(V)$.*

Proof. By scaling V (by a^{-1}), we may assume that $a = 1$. Let us prove the last statement first. Suppose $1 + \delta \in Q(V)$ with $\delta \in \Delta(V)^{-1} \cdot \mathfrak{p}$. By 1.4, $V \cong F\mathcal{A}(1+\delta, \alpha)$ with $|\alpha| = |\Delta(V)|$. By Hensel's Lemma, $X^2 + X + \alpha\delta = 0$ has a solution $X = \xi$ in F. Putting $\eta = \xi\alpha^{-1}$, we have

$$1 = (1+\delta) + \eta + \eta^2\alpha \in Q(V).$$

Now the first part. Since $\Delta(V) \sim \pi$, it is clear that $Q(V)$ contains elements whose orders are of either parity, and therefore also elements of the form $1 + d$ where $d \in \mathfrak{p}$. If $1 \notin Q(V)$, we can choose such an element $1 + \delta$ with $\mathfrak{d}(1+\delta) = \delta\mathfrak{o}$ minimal (by the first part of the proof); in particular $\delta \not\sim 1$. Suppose that $\delta \notin \Delta(V)^{-1} \cdot \mathfrak{o}$. Choose β so that $\beta^2\alpha \equiv \delta \bmod \delta\mathfrak{p}$ (where α is defined above). It is easy to see that

$$1 + \delta' = (1+\delta) + \beta + \beta^2\alpha \in Q(V)$$

has $\mathfrak{d}(1+\delta') \subset \delta\mathfrak{o}$, contradicting the choice of $1+\delta$. Therefore $\delta \in \Delta(V)^{-1} \cdot \mathfrak{o}$ as required.

LEMMA 1.22. *Let L be a binary unimodular lattice with $\Delta(L) \notin \mathfrak{o}$ and let a be a norm generator of $\mathfrak{g}L$ which is not represented by FL, i.e. $a \notin Q(FL)$. Then for a suitable β,*

$$L \cong \mathcal{B}(a + \rho\beta^{-1}, \beta).$$

Proof. Suppose $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. We have $L \cong \mathcal{B}(a', \beta')$ for some norm generator a' of $\mathfrak{g}L$. Thus $a' = \epsilon^2 a + \lambda$ where $\epsilon \in \mathfrak{u}$ and $\lambda \in b\mathfrak{o}$; in fact we may assume that $\mathfrak{d}(aa') = \lambda a\mathfrak{o}$. By replacing a' by $\epsilon^{-2}a'$ and β' by $\epsilon^2\beta'$, we may also assume that $\epsilon = 1$. By Lemma 1.21, there exists $c \in a\Delta(L)^{-1}\mathfrak{u}$ such that $a + c \in Q(FL)$ and moreover $|\lambda| \geqq |c|$. Thus $a + c \in \mathfrak{g}L$ and it follows by equality of norm groups (Sah's theorem) that $L \cong \mathcal{B}(a+c, \beta)$ where $\beta = \Delta(L) \cdot (a+c)^{-1}$. Now $c\beta \in \mathfrak{u}$; if $c\beta \in \wp(F)$, then $a\beta \equiv \Delta(L) \bmod \wp(F)$ whence $FL \cong \mathcal{A}(\beta, a)$ by 1.4—a contradiction since $a \notin Q(FL)$. Thus $c\beta \equiv \rho \bmod \wp(F)$. Thus $F\mathcal{B}(a + \rho\beta^{-1}, \beta) \cong F\mathcal{B}(a + c, \beta)$ since they have the same Arf Invariant, and β is represented by both. Since $\mathfrak{g}\mathcal{B}(a + \rho\beta^{-1}, \beta) = \mathfrak{g}L$, $L \cong \mathcal{B}(a + \rho\beta^{-1}, \beta)$ as required.

LEMMA 1.23. *Let L be a non-defective modular lattice with* $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. *Then* $\Delta(L) \in ab(\mathfrak{s}L)^{-2} + \mathfrak{o}$.

Proof. We may assume L to be unimodular. If L is binary, then the result follows from Lemma 1.20. The general result is now a consequence of the fact that L has a quasi-diagonal splitting, and also $a'\mathfrak{o}^2 + b'\mathfrak{o} \subseteq \mathfrak{g}L \Rightarrow a'b'\mathfrak{o} \subseteq ab\mathfrak{o}$.

LEMMA 1.24. *Let V be a quaternary non-defective space, and* $\mathfrak{g} = a\mathfrak{o}^2 + b\mathfrak{o} \supset \mathfrak{o}$ *a norm group such that* $\Delta(V) \in ab\mathfrak{o}$. *Then there is a unimodular lattice*

$$L \cong \mathcal{A}(b, \omega b^{-1}) \perp \mathcal{B}(a, b'), \quad \omega = 0 \text{ or } \rho, \tag{2}$$

on V with $\mathfrak{g}L = \mathfrak{g}$.

Proof. Let $\delta = \Delta(V)$. It is easy to show by means of the rules given at the bottom of p. 813 in [8], that the spaces

$$V_1 \cong F\mathcal{A}(b, 0) \perp F\mathcal{B}(a, \delta a^{-1}),$$
$$V_2 \cong F\mathcal{A}(b, \rho b^{-1}) \perp F\mathcal{B}(a, (\delta + \rho)a^{-1})$$

have non-isomorphic Clifford algebras; they are therefore not isometric. By Arf's theorem (p. 167 in [1]), $V \cong V_1$ or $V \cong V_2$. In the first case put $\omega = 0$ and $b' = \delta a^{-1}$; in the second, put $\omega = \rho$ and $b' = (\delta + \rho)a^{-1}$. Then (2) holds by Sah's theorem on integral equivalence (see 1.15), since $\delta \in ab\mathfrak{o} \Rightarrow b' \in b\mathfrak{o}$.

COROLLARY 1.25. *Let K be a unimodular lattice of dimension $\geqq 4$, and let $\mathfrak{g}$ be a norm group containing $\mathfrak{g}K$. Then there is a unimodular lattice J on FK with* $\mathfrak{g}J = \mathfrak{g}$.

Proof. We may suppose that $\mathfrak{g}K \subset \mathfrak{g}$, so that in particular, $\mathfrak{s}K \subset \mathfrak{g}$. By Corollary 1.10, $FK = H \perp V$ where H is a hyperbolic space, and V is quaternary. Since $\Delta(K) = \Delta(V)$, Lemmas 1.23 and 1.24 show the existence of a unimodular lattice L on V with $\mathfrak{g}L = \mathfrak{g}$. Let h be a unimodular hyperbolic lattice on H (or 0 if $H = 0$); then $J = h \perp L$ is the required lattice.

LEMMA 1.26. *Let L be a quaternary unimodular lattice with* $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. *Then*

$$L \cong \mathcal{A}(b, \omega b^{-1}) \perp \mathcal{B}(a, b'), \quad \omega = 0 \text{ or } \rho;$$

if $b \in \mathfrak{p}$, then ω must $= 0$ (and so $\mathcal{A}(b, \omega b^{-1}) \cong \mathcal{H}$).

Proof. If $\mathfrak{o} \subset \mathfrak{g}L$, this is a consequence of Lemma 1.24 and Sah's

theorem on integral equivalence. If $\mathfrak{g}L = \mathfrak{o}$, the result follows from Lemma 4.5 in [8], and an application of Lemma 1.6 if necessary (to show that a can be chosen as any norm generator).

Since $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o} = a\mathfrak{o}^2 + b\mathfrak{o}^2$, and by Lemma 1.16(i), it is easy to deduce:

COROLLARY. *If L is an $\mathfrak{a}$-modular lattice of dimension $\geqq 4$, then* $\mathfrak{g}L = Q(L)$.

1.27. The lattice $L^{\mathfrak{a}}$. Let $\mathfrak{a}$ be any ideal. Define

$$L^{\mathfrak{a}} = \{x \in L \mid B(x, L) \subseteq \mathfrak{a}\}.$$

If $\mathfrak{a} = 0$, then clearly $L^{\mathfrak{a}} = L^*$. If $\mathfrak{a} \neq 0$, $L^{\mathfrak{a}} \subseteq L$ is a lattice on FL, i.e. $F(L^{\mathfrak{a}}) = FL$. If L is $\mathfrak{b}$-modular, $(\mathfrak{b} \neq 0)$, then $L^{\mathfrak{a}} = L$ if $\mathfrak{b} \subseteq \mathfrak{a}$, otherwise $L^{\mathfrak{a}} = \mathfrak{a}\mathfrak{b}^{-1}L$. Also $(L \perp K)^{\mathfrak{a}} = L^{\mathfrak{a}} \perp K^{\mathfrak{a}}$. Thus if $L = L_1 \perp \cdots \perp L_T \perp L^*$ is a Jordan splitting of L with $\mathfrak{s}L_i = \mathfrak{b}_i$, then

$$L^{\mathfrak{a}} = \mathfrak{a}\mathfrak{b}_1^{-1}L_1 \perp \cdots \perp \mathfrak{a}\mathfrak{b}_{r-1}^{-1}L_{r-1} \perp L_r \perp \cdots \perp L_T \perp L^*$$

if $\mathfrak{b}_{r-1} \supset \mathfrak{a} \supseteq \mathfrak{b}_r$. Hence if $\mathfrak{a} = \mathfrak{b}_r$, $\mathfrak{g}L_r \subseteq \mathfrak{g}L^{\mathfrak{a}}$.

LEMMA 1.28. *Let $L = L_1 \perp \cdots \perp L_T \perp L^*$ be a Jordan splitting, and let I be any subset of $\{i \mid \dim L_i \geqq 4\}$. Let $\mathfrak{s}L_i = \mathfrak{s}(i)$, $1 \leqq i \leqq T$. Then there is another Jordan splitting $L = L_1' \perp \cdots \perp L_T' \perp L^*$ such that*

$$\begin{aligned} \mathfrak{g}L_i' &= \mathfrak{g}L^{\mathfrak{s}(i)} && \text{if } i \in I \\ L_i' &\cong L_i && \text{if } i \notin I. \end{aligned}$$

Proof. If $i \in I$, there is an $\mathfrak{s}(i)$-modular lattice L_i'' on FL_i with $\mathfrak{g}L_i'' = \mathfrak{g}L^{\mathfrak{s}(i)}$. Define

$$L' = (\underset{i \in I}{\perp} L'') \perp (\underset{i \notin I}{\perp} L) \perp L^*,$$

and let $L' = L_1' \perp \cdots \perp L_T' \perp L^*$ be the Jordan splitting naturally induced on L'. It is a straightforward matter to check that $L' \cong L$ by Theorem 5.5 in [8] since L and L' have the same invariants used in that theorem, and also $FL_i' \cong FL_i$ by the definition of L'. And it is obvious how this leads to the required Jordan splitting of L.

1.29. The dual lattice $L^\#$. Let L be a non-defective lattice. Define

$$L^\# = \{x \in FL \mid B(x, L) \subseteq \mathfrak{o}\}.$$

This dual lattice is not defined for a defective lattice. It is easy to check that $L^\#$ is a lattice on FL, $L^{\#\#} = L$, $L^\# = (\mathfrak{s}L)^{-1}L$ if L is modular, and that

$(L \perp K)^{\#} = L^{\#} \perp K^{\#}$. Thus if $L = L_1 \perp \cdots \perp L_T$ is a Jordan splitting, then $L^{\#} = (\mathfrak{s}L_T)^{-1}L_T \perp \cdots \perp (\mathfrak{s}L_1)^{-1}L_1$ is also a Jordan splitting.

PRINCIPLE OF DUALITY. *If l and L are non-defective lattices of the same dimension, then $l \to L$ if and only if $L^{\#} \to l^{\#}$.*

Proof. Suppose $l \subseteq L$. It follows directly from the definition that $L^{\#} \subseteq l^{\#}$. Thus $l \to L$ implies that $L^{\#} \to l^{\#}$. Hence $L^{\#} \to l^{\#}$ implies that $l^{\#\#} \to L^{\#\#}$, i.e. $l \to L$.

1.30. The lattice $L_{\mathfrak{h}}$. Let L be a non-defective lattice and $\mathfrak{h}$ a norm group with $\mathfrak{o}^2$-rank $= 2$. Then $L_{\mathfrak{h}}$ is the lattice generated by $\{x \in L \mid Q(x) \in \mathfrak{h}\}$. It is clearly a lattice on FL. The following relationships are easily verified.

$$L_{\mathfrak{h}} \subseteq L;\ \mathfrak{g}L \subseteq \mathfrak{h} \Rightarrow L_{\mathfrak{h}} = L.$$

$$\text{If } FK = FL,\ K \subseteq L \Rightarrow K_{\mathfrak{h}} \subseteq L_{\mathfrak{h}},\ L_{\mathfrak{h}} \subseteq K \subseteq L \Rightarrow K_{\mathfrak{h}} = L_{\mathfrak{h}}.$$

$$\mathfrak{h}' \subseteq \mathfrak{h} \Rightarrow (L_{\mathfrak{h}})_{\mathfrak{h}'} = L_{\mathfrak{h}'}.$$

$$L_{\mathfrak{h}} = J \perp K_{\mathfrak{h}} \text{ if } L = J \perp K \text{ and } \mathfrak{g}J \subseteq \mathfrak{h}.$$

$$L_{\mathfrak{h}} = L \text{ if } L \text{ is hyperbolic.}$$

The importance of this new lattice is due to the obvious fact that $l \to L \Longleftrightarrow l \to L_{\mathfrak{h}}$ if $\mathfrak{g}l \subseteq \mathfrak{h}$. Therefore we must be able to calculate $L_{\mathfrak{h}}$. This is usually an easy matter when $L = J \perp K$ with $\mathfrak{g}J \subseteq \mathfrak{h}$ (whence $L_{\mathfrak{h}} = J \perp K_{\mathfrak{h}}$) and $\dim K = 2$. The determination of $L_{\mathfrak{h}}$ is often made simpler by scaling by some suitable $\alpha \neq 0$, and using the formulas

$$L_{\mathfrak{h}} = (L^{\alpha})_{\alpha\mathfrak{h}},\ (L^{\#})_{\mathfrak{h}} = \alpha(((L^{\alpha})^{\#})_{\alpha^{-1}\mathfrak{h}}).$$

LEMMA 1.31. *Let K be a binary non-defective lattice and $\mathfrak{h}$ a norm group with $\mathfrak{o}^2$-rank $= 2$. Then either $\mathfrak{n}K_{\mathfrak{h}} \subseteq \mathfrak{h}\mathfrak{o}$ or $K_{\mathfrak{h}}$ is hyperbolic.*

Proof. Since $\mathfrak{n}K_{\mathfrak{h}} \subseteq \mathfrak{n}K_{\mathfrak{h}\mathfrak{o}}$ and $(\mathcal{H}_j)_{\mathfrak{h}} = \mathcal{H}_j$, it suffices to prove the lemma for $\mathfrak{h}$ an ideal of $\mathfrak{o}$. Suppose $\mathfrak{h} \subset \mathfrak{n}K_{\mathfrak{h}}$. Then there are vectors $x_1, \cdots, x_n$ in K such that $Q(x_1), \cdots, Q(x_n) \in \mathfrak{h}$—in particular $\sum x_i \in K_{\mathfrak{h}}$—and such that $\mathfrak{n}K_{\mathfrak{h}} = Q(\sum x_i)\mathfrak{o}$. For at least two of them, say x_1 and x_2, $\mathfrak{n}K_{\mathfrak{h}} \subseteq B(x_1, x_2)\mathfrak{o}$. Define $k = \mathfrak{o}x_1 + \mathfrak{o}x_2 \subseteq K_{\mathfrak{h}}$. Since $\mathfrak{s}K_{\mathfrak{h}} \subseteq \mathfrak{s}k$, we must have $k = K_{\mathfrak{h}}$, whence $\mathfrak{n}K_{\mathfrak{h}} = \mathfrak{s}K_{\mathfrak{h}}$. Also

$$|\Delta(K_{\mathfrak{h}})| \leqq |Q(x_1)Q(x_2)B(x_1, x_2)^{-2}| < 1$$

which implies that $\Delta(K_{\mathfrak{h}}) = 0$. Therefore $K_{\mathfrak{h}}$ is hyperbolic by Lemma 1.16(iii).

1.32. Determination of $K_{\mathfrak{h}}$ when K is binary unimodular. First let

us consider the case when $\mathfrak{h}$ is an ideal. Write $K \cong \mathcal{B}(\alpha, \beta)$. If $\Delta(K) = 0$, then $K_{\mathfrak{h}} \cong \mathcal{B}_i(\pi^{2i}\alpha, 0)$ where i is the smallest integer $\geqq 0$ such that $\pi^{2i}\alpha \in \mathfrak{p}^i + \mathfrak{h}$. This is easily shown using "domination," as is the next case. Namely if $\Delta(K) \neq 0$, then $K_{\mathfrak{h}} \cong \mathcal{A}_{i+j}(\pi^{2i}\alpha, \pi^{2j}\beta)$ where i (resp. j) is the smallest integer $\geqq 0$ such that $\pi^{2i}\alpha \in \mathfrak{h}$ (resp. $\pi^{2j}\beta \in \mathfrak{h}$).

Now suppose that $\mathfrak{h} = a\mathfrak{o}^2 + b\mathfrak{o}$ is any norm group with $ab \neq 0$. By the above case and by Lemma 1.31, we can assume that $\mathfrak{n}K \subseteq \mathfrak{h}\mathfrak{o} = a\mathfrak{o}$. To enumerate all possible cases would be too time-consuming and repetitious. We shall therefore confine ourselves to a few typical cases, some of which are needed later in the paper. The proofs involve straightforward applications of domination, and will therefore only be sketched as briefly as possible. If $K \cong \mathcal{B}(\alpha, \beta)$, we can write $\alpha = \xi^2 a + \eta$ (resp. $\beta = \theta^2 a + \zeta$) where $\mathfrak{d}(a\alpha) = a\eta\mathfrak{o}$ (resp. $\mathfrak{d}(a\beta) = a\zeta\mathfrak{o}$) and $\alpha = \eta$ if $\alpha \not\sim a$ (resp. $\beta = \zeta$ if $\beta \not\sim a$).

1.) If $\Delta(K) = 0$, then $K_{\mathfrak{h}} \cong \mathcal{B}_i(\pi^{2i}\alpha, 0)$ where i is defined as the smallest integer $\geqq 0$ such that $\pi^{2i}\eta \in \mathfrak{p}^i + b\mathfrak{o}$. Indeed, it is clear that $\mathcal{B}_i(\pi^{2i}\alpha, 0) \rightarrow K_{\mathfrak{h}}$, since $\mathcal{B}_i(\pi^{2i}\alpha, 0) \cong \mathcal{B}_i(\xi^2\pi^{2i}a, 0)$ by equality of norm groups if $\pi^{2i}\eta \in \mathfrak{p}^i$. On the other hand suppose $x \in K$ and $Q(x) \in \mathfrak{h}$. Then

$$Q(x) = c^2(\xi^2 a + \eta) + cd = e^2 a + b'$$

where $c, d, e \in \mathfrak{o}$ and $b' \in b\mathfrak{o}$. Now multiply the equation by a and compare quadratic defects. The result is $c^2\eta \in c\mathfrak{o} + b\mathfrak{o}$, whence $K_{\mathfrak{h}} \rightarrow \mathcal{B}_i(\pi^{2i}\alpha, 0)$, and the result follows. This proof is typical of those required for calculating $K_{\mathfrak{h}}$.

2.) If $\Delta(K) = \rho$, then $K_{\mathfrak{h}} = K_{b\mathfrak{o}}$ if $FK \cong F\mathcal{A}(b, \rho b^{-1})$, otherwise (if $FK \cong F\mathcal{A}(a, \rho a^{-1})$) we have $K_{\mathfrak{h}} \cong \mathcal{B}_i(\pi^{2i}\alpha, \beta)$ where i is the smallest integer $\geqq 0$ such that $\pi^{2i}\eta \in \mathfrak{p}^i + b\mathfrak{o}$.

3.) Suppose $\Delta(K) \notin \mathfrak{o}$. We shall consider the case $\alpha \sim a$, $\beta \sim b$. Let i be the smallest integer $\geqq 0$ such that $\pi^{2i}\eta \in \mathfrak{p}^i + \beta\mathfrak{o} + b\mathfrak{o}$.

First suppose $a \in Q(FK)$. If $\beta \in b\mathfrak{o}$, $K_{\mathfrak{h}} \cong \mathcal{B}_i(\pi^{2i}\alpha, \beta)$. Otherwise (if $\beta \notin b\mathfrak{o}$), $\mathcal{B}_i(\pi^{2i}\alpha, \beta) \cong \mathcal{B}_i(\pi^{2i}\xi^2 a, \beta')$ for suitable $\xi \in \mathfrak{o}$, and we let j be the smallest integer $\geqq 0$ such that $\pi^{2j}\beta' \in \mathfrak{p}^{i+j} + b\mathfrak{o}$. Then $K_{\mathfrak{h}} \cong \mathcal{B}_{i+j}(\pi^{2i}\xi^2 a, \pi^{2j}\beta')$.

Now suppose that $a \notin Q(FK)$. Define $L \cong \mathcal{B}_i(\pi^{2i}\alpha, \beta)$. If $\beta \in b\mathfrak{o}$, then $K_{\mathfrak{h}} \cong L$ as before. Suppose that $\beta \notin b\mathfrak{o}$. It is easy to see that $K_{\mathfrak{h}} \subseteq L \subseteq K$ (after suitable identifications have been made), and therefore $K_{\mathfrak{h}} = L_{\mathfrak{h}}$. Now $\pi^{2i}a$ is a norm generator of $\mathfrak{g}L$, whence

$$L \cong \mathcal{B}_i(\pi^{2i}(a + \rho\beta_1^{-1}), \beta_1)$$

for some $\beta_1 \in \beta\mathfrak{u}$, by Lemma 1.22. It follows from this that

$$K_{\mathfrak{h}} \cong \mathcal{B}_{i+j+k}(\pi^{2i+2k}(a + \rho\beta_1^{-1}), \pi^{2j}\beta_1) \tag{3}$$

where k (resp. j) is the smallest integer $\geqq 0$ such that $\pi^{2i+2k}\rho\beta_1^{-1} \in \mathfrak{b}\mathfrak{o}$ (resp. $\pi^{2j}\beta_1 \in \mathfrak{b}\mathfrak{o}$). This choice of k and j is dictated by the following considerations. An element $Q(x)$ represented by L is of the form $Q(x) = c^2\pi^{2i}a + Q'(x)$ where

$$Q'(x) = c^2\pi^{2i}\rho\beta_1^{-1} + cd\pi^i + d^2\beta_1,\ c \text{ and } d \in \mathfrak{o}.$$

By domination, $|Q'(x)| = \sup\{|c^2\pi^{2i}\rho\beta_1^{-1}|, |d^2\beta_1|\}$ and therefore $Q(x) \in \mathfrak{h} \Longleftrightarrow |Q'(x)| \leqq |b|$.

The expressions given above for $K_{\mathfrak{h}}$ are by no means easy to derive explicitly from a given expression $K \cong \mathscr{B}(\alpha,\beta)$; for example, it would be exceedingly difficult to derive the element β_1 of the last example, given α and β. However for the applications in which we are interested, it is always enough to be able to calculate $\mathfrak{g}K_{\mathfrak{h}}$, and this can be done directly from our formulas; e.g. it is enough to know $|\beta_1|$ in order to calculate $\mathfrak{g}K_{\mathfrak{h}}$ in (3).

§ 2. Theory of Maximal Lattices.

2. 1. *Definition.* Let $\mathfrak{a}$ be a non-zero ideal. A lattice L on a space V is said to be $\mathfrak{a}$-maximal on V if $\mathfrak{n}L \subseteq \mathfrak{a}$, and if $L \subset K \Rightarrow \mathfrak{a} \subset \mathfrak{n}K$.

LEMMA 2. 2. *An $\mathfrak{a}$-maximal lattice which is isotropic is split by an $\mathfrak{a}$-modular hyperbolic plane.*

Proof. Let x be a maximal isotropic vector in L. Then $B(x,L) = \mathfrak{a}$ since otherwise $\mathfrak{n}(\mathfrak{o}(\pi^{-1}x) + L) \subseteq \mathfrak{a}$. Choose $y \in L$ with $B(x,y)\mathfrak{o} = \mathfrak{a}$. Then $J = \mathfrak{o}x + \mathfrak{o}y$ is $\mathfrak{a}$-modular and therefore splits L by Lemma 1.17, since $\mathfrak{s}L \subseteq \mathfrak{n}L \subseteq \mathfrak{a}$. Clearly $\Delta(J) = 0$, and so J is an $\mathfrak{a}$-modular hyperbolic plane by 1.16(iii).

LEMMA 2. 3. *If L is a non-isotropic $\mathfrak{a}$-maximal lattice, then*

$$L = \{x \in FL \mid Q(x) \in \mathfrak{a}\}.$$

Thus two $\mathfrak{a}$-maximal lattices on the same non-isotropic space are equal.

Proof. Put $A = \{x \in FL \mid Q(x) \in \mathfrak{a}\}$ and let x and y be any vectors in A. If $B(x,y) \notin \mathfrak{a}$, then the space $U = Fx + Fy$ has $\Delta(U) = 0$, and is therefore a hyperbolic plane (cf. 1.3). This is a contradiction, and so $B(x,y) \in \mathfrak{a}$, whence $x + y \in A$. Thus A is an $\mathfrak{o}$-module. Clearly $L \subseteq A$; if $L \subset A$, then $L \subset L + \mathfrak{o}z \subseteq A$ where z is any vector in A not in L, and this contradicts the maximality of L. Therefore $L = A$ as required.

THEOREM 2. 4. *Two $\mathfrak{a}$-maximal lattices on the same space are isometric.*

Proof. Let L and L' be the two lattices. By Lemma 2.2, $L = K \perp H$ and $L' = K' \perp H'$ where K and K' are non-isotropic $\mathfrak{a}$-maximal lattices, and H and H' are $\mathfrak{a}$-modular hyperbolic lattices (or zero)—note that any component of an $\mathfrak{a}$-maximal lattice is $\mathfrak{a}$-maximal. By the Witt-Arf theorem (cf. 1.4), $FK \cong FK'$, $FH \cong FH'$. Thus trivially, $H \cong H'$. By Lemma 2.3, the isometry mapping FK onto FK' must map K onto K', whence $K \cong K'$. Therefore $L \cong L'$ as required.

2.5. The volume of a lattice. Let $L = \mathfrak{o}x_1 + \cdots + \mathfrak{o}x_n$ be a non-defective lattice. The determinant $d_B(x_1, \cdots, x_n) = \det(B(x_i, x_j))$ is called the *discriminant of B in the basis $x_1, \cdots, x_n$.* If another basis $y_1, \cdots, y_n$ is chosen for L, it is well-known that $d_B(y_1, \cdots, y_n) = \epsilon^2 d_B(x_1, \cdots, x_n)$ for some unit ϵ. Thus in particular, the ideal

$$\mathfrak{v}(L) = d_B(x_1, \cdots, x_n)\mathfrak{o}$$

is an invariant of L. It can be calculated from a Jordan splitting $L = L_1 \perp \cdots \perp L_T$ as follows:

$$\mathfrak{v}(L) = (\mathfrak{s}L_1)^{d(1)} \cdot (\mathfrak{s}L_2)^{d(2)} \cdot \cdots \cdot (\mathfrak{s}L_T)^{d(T)} \tag{4}$$

where $d(i) = \dim L_i$. This can be easily proved by taking a quasi-diagonal splitting for each component L_i.

Now suppose that L is a defective lattice, and consider two splittings $L = L^\ddagger \perp L^* = L^{\ddagger\ddagger} \perp L^*$. Any Jordan splitting of $L^\ddagger$ or $L^{\ddagger\ddagger}$ induces a Jordan splitting of L. By the invariance of the quantities T, $\dim L_i$, and $\mathfrak{s}L_i$ for Jordan splittings of L, it follows from (4) that $\mathfrak{v}(L^\ddagger) = \mathfrak{v}(L^{\ddagger\ddagger})$. It follows from this that $\mathfrak{v}(L)$ as defined below is well-defined.

Definition. Let L be a lattice, and suppose that $L = L^\ddagger \perp L^*$. Take a basis $x_1, \cdots, x_n$ for $L^\ddagger$, and suppose that $\mathfrak{g}L^* = a\mathfrak{o}^2 + b\mathfrak{o}$. The volume $\mathfrak{v}(L)$ is defined to be $\mathfrak{v}(L) = d_B(x_1, \cdots, x_n) \cdot \mathfrak{a}$ where

$$\mathfrak{a} = \begin{cases} \mathfrak{o} \text{ if } a = b = 0, \\ a\mathfrak{o} \text{ if } a \neq 0 = b, \\ ab\mathfrak{o} \text{ if } ab \neq 0. \end{cases}$$

LEMMA 2.6. *If L and J are two lattices on the same space, then $L \subset J$ implies that $\mathfrak{v}(L) \subset \mathfrak{v}(J)$.*

Proof. Write $L = L^\ddagger \perp L^*$, $J = J^\ddagger \perp J^*$. Clearly $L^* \subseteq J^*$. Define a map $\phi: L^\ddagger \to J^\ddagger$ by $x = \phi x + z$ where $z \in J^*$. Thus $B(x, y) = B(\phi x, \phi y)$ for all $x, y \in L^\ddagger$, whence ϕ is injective, and $\mathfrak{v}(L^\ddagger) = \mathfrak{v}(\phi L^\ddagger)$. By the elementary divisor theorem, there is a basis $x_1, \cdots, x_n$ for $J^\ddagger$ and integers $\alpha_1, \cdots, \alpha_n$ such that $\alpha_1 x_1, \cdots, \alpha_n x_n$ is a basis for $\phi L^\ddagger$; thus $\mathfrak{v}(L^\ddagger) = \mathfrak{v}(\phi L^\ddagger) \subseteq \mathfrak{v}(J^\ddagger)$.

Suppose that $L^* = J^*$. Then $\phi L^\dagger \subseteq L$ since $\phi x = x + z$, and so $\phi L^\dagger \subset J^\dagger$. Therefore $\mathfrak{d}(L^\dagger) = \mathfrak{d}(\phi L^\dagger) \subset \mathfrak{d}(J^\dagger)$ whence

$$\mathfrak{d}(L) = \mathfrak{d}(L^\dagger)\mathfrak{d}(L^*) \subset \mathfrak{d}(J^\dagger)\mathfrak{d}(J^*) = \mathfrak{d}(J).$$

It is easy to show that $L^* \subset J^*$ implies that $\mathfrak{g}L^* \subset \mathfrak{g}J^*$ (recall that if $\mathfrak{g}L^* = a\mathfrak{o}^2 + b\mathfrak{o}$ for example, then $L^* \cong \langle a \rangle \perp \langle b \rangle$). Thus $\mathfrak{d}(L^*) \subset \mathfrak{d}(J^*)$ whence $\mathfrak{d}(L) \subset \mathfrak{d}(J)$.

LEMMA 2.7. *Let L be a lattice in a space V and suppose that $\mathfrak{n}L \subseteq \mathfrak{a} \neq 0$. Then there is an $\mathfrak{a}$-maximal lattice on V which contains L.*

Proof. Let J be any lattice on V. Choose $\alpha \in \mathfrak{o}$, $\alpha \neq 0$, such that $B(L, \alpha J) \subseteq \mathfrak{a}$, $\mathfrak{n}(\alpha J) \subseteq \mathfrak{a}$. Then $L \subseteq L + \alpha J$ and $\mathfrak{n}(L + \alpha J) \subseteq \mathfrak{a}$. It suffices to find an $\mathfrak{a}$-maximal lattice containing $L + \alpha J$; we may therefore assume that $FL = V$.

If L is not $\mathfrak{a}$-maximal, there exists $L_1 \supset L$ such that $\mathfrak{n}L_1 \subseteq \mathfrak{a}$. If L_1 is not maximal, there exists $L_2 \supset L_1$ such that $\mathfrak{n}L_1 \subseteq \mathfrak{a}$. Continuing in this fashion, we get a chain of lattices $L \subset L_1 \subset L_2 \subset \cdots$ and a chain of ideals $\mathfrak{d}(L) \subset \mathfrak{d}(L_1) \subset \mathfrak{d}(L_2) \subset \cdots$ by Lemma 2.6. But it is easy to see that $\mathfrak{n}L_i \subseteq \mathfrak{a} \Rightarrow \mathfrak{d}(L_i) \subseteq \mathfrak{a}^d$ where $d = \dim L$. Thus the chain of ideals breaks off at some point, and therefore so does the chain of lattices, i.e. there is an $\mathfrak{a}$-maximal lattice containing L.

COROLLARY 2.8. *Let V be any space and $\mathfrak{a}$ a non-zero ideal. Then V supports an $\mathfrak{a}$-maximal lattice, i.e. there is an $\mathfrak{a}$-maximal lattice on V.*

Proof. By Lemma 2.7, there is an $\mathfrak{a}$-maximal lattice on V which contains the trivial lattice 0.

The importance of maximal lattices in the theory of integral representation is due to the following lemma. Lemma 2.11 and its corollary are useful complements for it.

LEMMA 2.9. *Suppose that L contains an $\mathfrak{a}$-maximal lattice on FL (in particular L may itself be $\mathfrak{a}$-maximal). If l is another lattice with $\mathfrak{n}l \subseteq \mathfrak{a}$ and $Fl \to FL$, then $l \to L$.*

Proof. We may assume that l is in FL. Let $M \subseteq L$ be an $\mathfrak{a}$-maximal lattice on FL. By Lemma 2.7, there is an $\mathfrak{a}$-maximal lattice M' on FL containing l, and $M' \cong M$ by Theorem 2.4. Then $l \to M' \cong M \to L$ as required.

LEMMA 2.10. *If $0 \neq \mathfrak{b} \subseteq \mathfrak{a}$, an $\mathfrak{a}$-maximal lattice L on a space V contains some $\mathfrak{b}$-maximal lattice K on V.*

Proof. By Corollary 2.8, there is a $\mathfrak{b}$-maximal lattice K' on V. By Lemma 2.9, $K' \to L$; the image K of K' under this representation is the required lattice.

LEMMA 2.11. *Let L be a non-defective lattice and suppose that $0 \neq \mathfrak{a} \subseteq (\mathfrak{n}L^{\#})^{-1}$. Then there is an $\mathfrak{a}$-maximal lattice on FL contained in L.*

Proof. By Lemma 2.10, we may suppose that $\mathfrak{a} = (\mathfrak{n}L^{\#})^{-1}$. There is a $(\mathfrak{n}L^{\#})$-maximal lattice M on FL containing $L^{\#}$ by Lemma 2.7. The lattice $J = \mathfrak{a}M$ is $\mathfrak{a}$-maximal and

$$J \subseteq (\mathfrak{s}M)^{-1}M \subseteq M^{\#} \subseteq L$$

since $\mathfrak{s}M \subseteq \mathfrak{n}M = \mathfrak{a}^{-1}$.

COROLLARY 2.12. *A modular lattice L with $\mathfrak{n}L = \mathfrak{s}L$ is $\mathfrak{n}L$-maximal. In particular $\mathfrak{H}_i$ and $\mathfrak{A}_i(\pi^i\epsilon, \pi^i\rho\epsilon^{-1})$, $\epsilon \in \mathfrak{u}$, are $\mathfrak{p}^i$-maximal.*

§ 3. A Structure Theorem for Representations.

THEOREM 3.1. *Suppose that $l \to L$ where L is non-defective. Let $\mathfrak{a}$ be a non-zero ideal contained in $(\mathfrak{s}L^{\#})^{-1}$. Suppose $l = l_1 \perp l_2$ where l_1 is non-defective, $\mathfrak{s}l_2 \subseteq \mathfrak{a} \subseteq (\mathfrak{s}l_1 + \mathfrak{s}L^{\#})^{-1}$, $\dim l_1 < \dim L$. Then there exists an $\mathfrak{a}$-modular lattice k such that $l_2 \to k$, $l_1 \perp k \to L$, $\dim(l_1 \perp k) = \dim L$.*

Proof. We can assume that $l \subseteq L$. Write $FL = Fl_1 \perp V$ and let l' be any lattice on V. For suitably large m, $\mathfrak{p}^m l' \subseteq L$ and $B(l_2 + \mathfrak{p}^m l', \mathfrak{p}^m l') \subseteq \mathfrak{a}$. It is clear that replacing l_2 by $l_2 + \mathfrak{p}^m l'$ allows us to assume that $\dim l = \dim L$. In particular $l_2 \neq 0$.

Suppose that l_2 is not $\mathfrak{a}$-modular. Then there is a maximal vector $x \in l_2$ such that $B(x, l_2) \subset \mathfrak{a}$. If x *is not* a maximal vector *in* L, then $\pi^{-1}x \in L$ and we set $J = \mathfrak{o}(\pi^{-1}x) + l_2$. Then $\mathfrak{s}J \subseteq \mathfrak{a}$. If x *is* a maximal vector *in* L, there is a vector $y \in L$ such that $B(x, y)\mathfrak{o} = B(y, L) = \mathfrak{a}$ (consider a quasi-diagonal splitting of L). Let $y = y_1 + y_2$ where $y_1 \in Fl_1$, $y_2 \in Fl_2$. Then $B(y_1, l_1) = B(y, l_1) \subseteq B(y, L) = \mathfrak{a}$, whence $y_1 \in l_1 \subseteq L$, and therefore $y_2 \in L$. Also $B(x, y_2)\mathfrak{o} = B(x, y)\mathfrak{o} = \mathfrak{a}$, so that $y_2 \notin l_2$ (since $B(x, l_2) \subset \mathfrak{a}$). Moreover if $J = \mathfrak{o}y_2 + l_2$, then $\mathfrak{s}J \subseteq \mathfrak{a}$.

Using the original vector x, we have found a lattice $J \supset l_2$ with $\mathfrak{s}J \subseteq \mathfrak{a}$. If $\dim l_2 = m$, $\mathfrak{v}(l_2) \subset \mathfrak{v}(J) \subseteq \mathfrak{a}^m$. If J is not $\mathfrak{a}$-modular, we repeat the process. Continuing in this manner, we get a chain of lattices on Fl_2 with volume at most $\mathfrak{a}^m$:

$$l_2 \subset J \subset J' \subset \cdots \subset L$$
$$\mathfrak{v}l_2 \subset \mathfrak{v}J \subset \mathfrak{v}J' \subset \cdots \subseteq \mathfrak{a}^m.$$

Thus the chain of lattices must break off at some time, i.e. there is an $\mathfrak{a}$-modular lattice k on Fl_2 such that $l_2 \subseteq k \subseteq L$.

LEMMA 3.2. *Let l and L be non-defective lattices with* $\dim l < \dim L$, $Fl \to FL$ *and* $\mathfrak{g}l \subseteq \mathfrak{g}L$. *Let $\mathfrak{a}$ be an ideal such that $\mathfrak{s}l^\#$ and $\mathfrak{s}L^\#$ are contained in $\mathfrak{a}^{-1}$. Then there is an $\mathfrak{a}$-modular lattice k such that $F(l \perp k) \cong FL$ and $\mathfrak{g}k \subseteq \mathfrak{g}L$.*

Proof. Let α and β be in $\mathfrak{g}L$ and $\mathfrak{p}^j \subseteq \mathfrak{s}l = \mathfrak{p}^s$. Then it is easy to see that

$$\langle\alpha\rangle \to \mathfrak{A}_s(\alpha, 0) \cong L', \mathfrak{A}_j(\alpha, \beta) \to \mathfrak{A}_s(\alpha, 0) \perp \mathfrak{A}_s(\beta, 0) \cong L'',$$

and that both $\mathfrak{g}L'$ and $\mathfrak{g}L''$ are in $\mathfrak{g}L$. Therefore there is a $\mathfrak{p}^s$-modular lattice M such that $l \to L \perp M$, $\mathfrak{g}M \subseteq \mathfrak{g}L$, and such that FM is a hyperbolic space. By Theorem 3.1, there is an $\mathfrak{a}$-modular lattice J such that $\mathfrak{g}J \subseteq \mathfrak{g}L$ and $F(l \perp J) \cong F(L \perp M)$. By 1.16(i), $J = J' \perp M'$ where $\dim J' \leqq 4$ and FM' is a hyperbolic space.

By removing the same number of binary lattices (each on a hyperbolic space) from M and M', we can achieve the following situation: there are lattices J', M' and M such that (i) J' and M' are $\mathfrak{a}$-modular and have norm groups contained in $\mathfrak{g}L$, (ii) FM and FM' are hyperbolic, and at least one of them is 0, (iii) $\dim J' \leqq 4$, and (iv) $F(l \perp J' \perp M') \cong F(L \perp M)$.

If $M = 0$, we may choose $k = J' \perp M'$. Suppose $M \neq 0$, so that by (ii), $M' = 0$. Since $\dim l < \dim L$, by (iii) and (iv) $\dim M = 2$, i.e. $FM \cong F\mathfrak{H}$, and $\dim J' = 4$. Now $FL \cong Fl \perp V$ for some space V, whence $Fl \perp V \perp F\mathfrak{H} \cong Fl \perp FJ'$, whence FJ' is isotropic. It is easy to see that J' can be split by some binary isotropic lattice h, and the lattice k determined by $J = h \perp k$ has the required properties.

§ 4. Representation Theory.

LEMMA 4.1. *Let l and L be $\mathfrak{a}$-modular lattices with* $\dim L - \dim l > 2$ *and* $\mathfrak{a} \neq 0$. *If* $\mathfrak{g}l \subseteq \mathfrak{g}L$, *there is an $\mathfrak{a}$-modular lattice k such that $l \perp k \cong L$ and* $\mathfrak{g}k = \mathfrak{g}L$.

Proof. By Theorem 1.11, $FL \cong Fl \perp V$ for some non-defective space of dimension $\geqq 4$. If $\mathfrak{g}L = \mathfrak{a}$, $l \to L$ by 2.12 and 2.9, and so $l \perp k' \cong L$ for some lattice k' by 1.17. By 1.25, there is an $\mathfrak{a}$-modular lattice k on Fk' with $\mathfrak{g}k = \mathfrak{g}L$. Then $l \perp k \cong L$ by Sah's equivalence theorem for modular lattices. Now suppose that $\mathfrak{a} \subset \mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$. Then $\Delta(V) = \Delta(l) + \Delta(L) \in ab\mathfrak{a}^{-2}$ by Lemma 1.23. We can suppose that $V = H \perp V'$ where H is hyper-

bolic and $\dim V' = 4$. Then $\Delta(V') = \Delta(V) \in ab\mathfrak{a}^{-2}$, so that there is an $\mathfrak{a}$-modular lattice k' on V' with $\mathfrak{g}k' = \mathfrak{g}L$ by Lemma 1.24. Let h be an $\mathfrak{a}$-modular hyperbolic lattice on H (or 0 if $H = 0$), and put $k = h \perp k'$. Then $l \perp k \cong L$ by Sah's theorem.

PROPOSITION 4.2. *Let l and L be non-defective lattices of the same dimension ≥ 3 and suppose that*

(5) $\qquad \Sigma_i(L) > \Sigma_i(l) + 2$ *when* $(\mathfrak{s}l^{\#})^{-1} \subset \mathfrak{p}^i \subseteq \mathfrak{s}L,$

(6) $\qquad \Sigma_i(L) \geqq \Sigma_i(l)$ *otherwise.*

Then $l \to L$ if and only if $Fl \cong FL$, $\mathfrak{g}l \subseteq \mathfrak{g}L$, and $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l^{\#}$.

Proof. The necessity is obvious. Note also that (6) is necessary (for all i) by Lemma 1.18; condition (5), on the other hand, is not necessary for representation.

We now pass on to the sufficiency. Condition (6) for $i = \operatorname{ord}(\mathfrak{s}L) - 1$ shows that $\mathfrak{s}l \subseteq \mathfrak{s}L$; $i = \operatorname{ord}(\mathfrak{s}L)$ in (5) implies that $\dim L_1 \geqq 4$ where L_1 is the first component in a Jordan splitting of L; if $(\mathfrak{s}l^{\#})^{-1} \subset \mathfrak{s}L$, $i = \operatorname{ord}(\mathfrak{s}l^{\#})^{-1} - 1$ in (5) implies that the last component l_t in a Jordan splitting of l has dimension $\geqq 4$; finally $i = \operatorname{ord}(\mathfrak{s}l^{\#})^{-1}$ in (6) implies that $\mathfrak{s}l_t \subseteq \mathfrak{s}L_T$ where L_T is the last component in a Jordan splitting of L.

The proof is by induction on the quantity

$$N(l, L) = (\operatorname{ord}(\mathfrak{s}l_t) - \operatorname{ord}(\mathfrak{s}L_1)) \cdot T \geqq 0.$$

Suppose $N(l, L) = 0$. This is equivalent to l and L being $(\mathfrak{s}L)$-modular. Then $\mathfrak{g}l \subseteq \mathfrak{g}L$ and $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l^{\#}$ imply that $\mathfrak{g}l = \mathfrak{g}L$, whence $l \cong L$ by Sah's theorem. We may therefore suppose that $N(l, L) > 0$ and that the theorem is true for all lattices l', L' satisfying the given conditions and for which $N(l', L') < N(l, L)$.

Let $L = L_1 \perp \cdots \perp L_T$ and $l = l_1 \perp \cdots \perp l_t$ be Jordan splittings; by Lemma 1.28, we may suppose that $\mathfrak{g}L_1 = \mathfrak{g}L$ and $\mathfrak{g}l_t^{\#} = \mathfrak{g}l^{\#}$.

First suppose that $\mathfrak{s}l_t = \mathfrak{s}L_T$. Condition (5) for $i = \operatorname{ord}(\mathfrak{s}l^{\#})^{-1} - 1$ shows that $\dim l_t - \dim L_T \geqq 4$. Since $\mathfrak{g}L_T^{\#} \subseteq \mathfrak{g}L^{\#} \subseteq \mathfrak{g}l^{\#} = \mathfrak{g}l_t^{\#}$, $l_t^{\#} \cong L_T^{\#} \perp k$ for some $\mathfrak{s}l_t^{\#}$-modular lattice k of dimension $\geqq 4$, with $\mathfrak{g}k = \mathfrak{g}l_t^{\#}$. Put $l_t' = k^{\#}$, so that $l_t \cong L_T \perp l_t'$ with $\mathfrak{g}l_t = \mathfrak{g}l_t'$, $\mathfrak{s}l_t = \mathfrak{s}l_t'$. Define

$$L' = L_T^{*} = L_1 \perp \cdots \perp L_{T-1} \text{ and } l' = l_t^{*} \perp l_t'.$$

By the Witt-Arf theorem, $FL' \cong Fl'$. Also $\mathfrak{g}l' \subseteq \mathfrak{g}l \subseteq \mathfrak{g}L = \mathfrak{g}L_1 \subseteq \mathfrak{g}L'$, and $\mathfrak{g}L'^{\#} \subseteq \mathfrak{g}L^{\#} \subseteq \mathfrak{g}l^{\#} = \mathfrak{g}l_t^{\#} \subseteq \mathfrak{g}l'^{\#}$. And $\Sigma_i(L') - \Sigma_i(l') = \Sigma_i(L) - \Sigma_i(l)$ for

all i while $\mathfrak{s}l'^{\#} = \mathfrak{s}l^{\#}$ and $\mathfrak{s}L' = \mathfrak{s}L$. Since $N(l', L') = \frac{T-1}{T} N(l, L) < N(l, L)$, $l' \to L'$ by the induction hypothesis, and therefore $l \cong l' \perp L_T \to L' \perp L_T \cong L$ as required.

From now on, we may suppose that $\mathfrak{s}l_t \subset \mathfrak{s}L_T$. Put $\mathfrak{h} = \mathfrak{g}L \cap \pi^{-2}\mathfrak{g}l_t$ and $\mathfrak{s}l_t = \mathfrak{p}^r$. We shall find a $\mathfrak{p}^{r-1}$-modular lattice k on Fl_t such that $l_t \to k$ and $\mathfrak{g}k = \mathfrak{h}$. Suppose for the moment that this has been done. It is a straightforward matter to check that $l' = l_t^* \perp k$ and L satisfy the conditions like (5) and (6). Trivially $\mathfrak{g}l' \subseteq \mathfrak{g}L$ and $Fl' \cong FL$. Now $\mathfrak{g}l'^{\#} = \mathfrak{g}l_t^{*\#} + \pi^{-2(r-1)}\mathfrak{h}$ and

$$\begin{aligned} \pi^{-2(r-1)}\mathfrak{h} &= (\pi^{-2(r-1)}\mathfrak{g}L) \cap \pi^{-2r}\mathfrak{g}l_t \\ &\supseteq \mathfrak{g}L^{\#} \cap \mathfrak{g}l_t^{\#} = \mathfrak{g}L^{\#}. \end{aligned}$$

Therefore $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l'^{\#}$. Since $N(l', L) < N(l, L)$, by the induction hypothesis $l' \to L$, and so $l = l_t^* \perp l_t \to l_t^* \perp k = l' \to L$ as required. It therefore suffices to find k. There are three cases to be considered: $\mathfrak{h}\mathfrak{o} = \pi^{-2}\mathfrak{n}l_t$, $\pi^{-1}\mathfrak{n}l_t$, or $\mathfrak{n}l_t$, since $\mathfrak{g}l_t \subseteq \mathfrak{h} \subseteq \pi^{-2}\mathfrak{g}l_t$.

By Lemma 3.2, there is a $\mathfrak{p}^{r-1}$-modular lattice K on Fl_t with $\mathfrak{g}K \subseteq \mathfrak{g}L$. This is the only fact we shall use in the proof outside of the existence of l_t itself and the definition of $\mathfrak{h}$. We may therefore assume that $\dim l_t = 4$. Indeed if $\dim l_t > 4$, then $l_t = h \perp l_t'$ where h is $\mathfrak{p}^r$-modular hyperbolic and $\dim l_t' = 4$, and $K = H \perp K'$ where H is $\mathfrak{p}^{r-1}$-modular hyperbolic and $\dim K' = 4$; thus $\mathfrak{g}K' \subseteq \mathfrak{g}L$ and $FK' \cong Fl_t'$. With the assumption that $\dim l_t = 4$, l_t has the form

$$l_t \cong \mathcal{A}_r(\beta, \omega\pi^{2r}\beta^{-1}) \perp \mathcal{B}_r(\alpha, \beta'), \quad \omega = 0 \text{ or } \rho,$$

where $\mathfrak{g}l_t = \alpha\mathfrak{o}^2 + \beta\mathfrak{o}$, and $\omega = 0$ if $\beta \in \mathfrak{p}^{r+1}$; this is a consequence of Lemma 1.26. Note that α and β can be any norm and base generators, respectively. Finally we may assume that $\mathfrak{g}L \supset \mathfrak{s}L$ if L is modular; in fact if $\mathfrak{g}L = \mathfrak{s}L$, then $l \to L$ by 2.9 and 2.12.

First Case. $\mathfrak{h}\mathfrak{o} = \pi^{-2}\mathfrak{n}l_t$. Since $\mathfrak{h} \subseteq \pi^{-2}\mathfrak{g}l_t$, we may suppose that α has been chosen so that $\pi^{-2}\alpha$ is a norm generator of $\mathfrak{h}$. Now either β or $\pi^{-2}\beta$ is a base generator of $\mathfrak{h}$. In the latter case we may choose

$$k \cong \mathcal{A}_{r-1}(\beta\pi^{-2}, \omega\pi^{2r}\beta^{-1}) \perp \mathcal{B}_{r-1}(\alpha\pi^{-2}, \beta').$$

Now suppose that β is a base generator of $\mathfrak{h}$; this can happen only when β is a base generator of $\mathfrak{g}L$ itself. In this case, we may choose

$$k \cong \mathcal{A}_{r-1}(\beta, \omega\pi^{2r-2}\beta^{-1}) \perp \mathcal{B}_{r-1}(\alpha\pi^{-2}, \beta')$$

unless $\omega\pi^{2r-2}\beta^{-1} \notin \mathfrak{h}$; this will happen when $|\beta| = |\pi^r|$, $\omega = \rho$, and (since β is a base generator of $\mathfrak{g}L$) when L is $\mathfrak{p}^{r-1}$-modular. By Lemma 1.26,

$$K \cong \mathcal{H}_{r-1} \perp \mathcal{B}_{r-1}(a', b').$$

If $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$, $a' = \xi^2 a + \eta$ where $\xi \in \mathfrak{o}$ and $\eta \in b\mathfrak{o} = \mathfrak{p}^r$. If $b' = 0$,

$$K \cong \mathcal{H}_{r-1} \perp \mathcal{B}_{r-1}(\xi^2 a, b'') \tag{7}$$

where $b'' = 0$, by equality of norm groups. If $b' \neq 0$, $\xi^2 a \in Q(\mathcal{B}_{r-1}(a', b'))$ by 1.6 or 1.21, and we can again put K in the form (7). Since $|\beta| = |b|$, $\lambda^2 a$ is a norm generator of $\mathfrak{g}l_t$ if $|\lambda^2 a| = |\alpha|$, and $(\lambda\pi^{-1})^2 a$ is a norm generator of $\mathfrak{h}$. Thus by equality of norm groups, $l_t \cong \mathcal{H}_r \perp \mathcal{B}_r(\lambda^2 a, \beta'')$ and we can choose

$$k \cong \mathcal{H}_{r-1} \perp \mathcal{B}_{r-1}((\lambda\pi^{-1})^2 a, \beta'').$$

Second Case. $\mathfrak{h}\mathfrak{o} = \pi^{-1}\mathfrak{n}l_t$. Then α must be a base generator of $\mathfrak{g}L$, and also of $\mathfrak{h}$. And $\beta\pi^{-2}$ is a norm generator of $\mathfrak{h}$. We can choose

$$k \cong \mathcal{A}_{r-1}(\beta\pi^{-2}, \omega\pi^{2r}\beta^{-1}) \perp \mathcal{B}_{r-1}(\alpha, \beta'\pi^{-2}) \tag{8}$$

except possibly in one case, when $\beta'\pi^{-2} \notin \mathfrak{h}$. It is easy to see that this could only happen if $|\alpha| = |\pi^r|$ and $\beta' = \rho\pi^{2r}\alpha^{-1}$, and so L is $\mathfrak{p}^{r-1}$-modular as in the previous case. But this situation cannot arise. Indeed $\beta \in \mathfrak{p}^{r+1} \Rightarrow \omega = 0$, whence K is isotropic and so $K \cong H \perp J$ where H is isotropic. By the Witt-Arf theorem, $FJ \cong F\mathcal{B}_r(\alpha, \rho\pi^{2r}\alpha^{-1})$, and it is easy to check that FJ cannot support a $\mathfrak{p}^{r-1}$-modular lattice with norm group in $\mathfrak{g}L$.

Third Case. $\mathfrak{h}\mathfrak{o} = \mathfrak{n}l_t$. Then α is a norm generator of $\mathfrak{h}$, and either β or $\pi^{-2}\beta$ is a base generator. In the former case $\mathfrak{g}l_t = \mathfrak{h} = \mathfrak{g}L$, and the existence of K and Corollary 1.25 show the existence of a lattice k on Fl_t with $\mathfrak{g}k = \mathfrak{h}$. By Lemma 1.26,

$$k \cong \mathcal{A}_{r-1}(\beta, \omega'\pi^{2(r-1)}\beta^{-1}) \perp \mathcal{B}_{r-1}(\alpha, \beta''), \ \omega' = 0 \text{ or } \rho.$$

Thus $l_t \cong \mathcal{A}_r(\beta, \omega'\pi^{2r}\beta^{-1}) \perp \mathcal{B}_r(\alpha, \beta''\pi^2)$ by equality of norm groups, whence $l_t \to k$. Now suppose $\beta\pi^{-2}$ is a base generator of $\mathfrak{h}$. Then we define k as in (8). This is valid since $|\alpha| > |\beta\pi^{-2}|$ implies that $\beta'\pi^{-2} \in \mathfrak{h}$—note in particular that $\alpha \notin \mathfrak{p}^r$. This finishes the proof of Proposition 4.2.

4.3. We want to relax condition (5) of Prop. 4.2. The more general case to be considered (Theorem 4.8) is

$$l^* = L^* = 0, \ \dim l = \dim L, \ \mathfrak{s}l \subseteq \mathfrak{s}L,$$

and, when $\dim L > 2$,

$$\begin{aligned} &\Sigma_i(L) > \Sigma_i(l) + 2 && \text{if } \mathfrak{q}^{-1} \subset \mathfrak{p}^i \subseteq \mathfrak{s}L \\ &\Sigma_i(L) > \Sigma_i(l) && \text{if } (\mathfrak{s}l^\#)^{-1} \subset \mathfrak{p}^i \subseteq \mathfrak{q}^{-1} \\ &\Sigma_i(L) \geqq \Sigma_i(l) && \text{otherwise} \end{aligned}$$

where, if $l = l_1 \perp \cdots \perp l_t$ is a Jordan splitting, $\mathfrak{q}^{-1} = \mathfrak{s}l_t$ if $\dim l_t > 2$, $\mathfrak{q}^{-1} = \mathfrak{s}l_{t-1}$ if $\dim l_t = 2$. Whenever l and L satisfy these conditions, we shall say that l *is well below* L. Suppose that this is the case, that $\dim L > 2$, and let $L = L_1 \perp \cdots \perp L_T$ be a Jordan splitting; then it is easy to check that

$$\mathfrak{q}^{-1} \subseteq \mathfrak{s}L_T,\ \dim L_1 > 2.$$

Now suppose that l is well below L, and that k is a non-defective lattice with $\dim k = \dim l$ and $\Sigma_i(k) \leqq \Sigma_i(l)$ for all i, (by Lemma 1.18, this is the case if $k \to l$). Then k is well below L. We leave this to the reader to check; one simply uses the definition and the relationships $(\mathfrak{s}k^\#)^{-1} \subseteq (\mathfrak{s}l^\#)^{-1}$ and $\mathfrak{r}^{-1} \subseteq \mathfrak{q}^{-1}$ where $\mathfrak{r}$ is the ideal for k analogous to $\mathfrak{q}$ for l.

Suppose that $\dim l < \dim L$; we say that l is well below L if there is a lattice J which is well below L, of the same dimension as L, and such that $\Sigma_i(l) \leqq \Sigma_i(J)$ for all i. Thus l may be defective, although L is not. Suppose $l = l^\natural \perp l^*$. It is easy to see that there is a modular lattice l' such that $\dim(l^\natural \perp l') = \dim L$ and $\Sigma_i(l) \leqq \Sigma_i(l^\natural \perp l') \leqq \Sigma_i(J)$ for all i. In particular $l^\natural \perp l'$ is well below L. If k is any modular lattice on Fl' with $\mathfrak{s}k = (\mathfrak{s}L^\# + \mathfrak{s}l^{\natural\#})^{-1}$, it is clear that $l^\natural \perp k$ is also well below L. Since $\Sigma_i(l^\natural \perp k) = \Sigma_i(l)$ if $\mathfrak{s}k \subset \mathfrak{p}^i$, it follows that, when $2 < \dim L > \dim l$, l is well below L if and only if

$$\text{(9)}\qquad \begin{aligned} &\Sigma_i(L) > \Sigma_i(l) + 2 && \text{if } (\mathfrak{s}L^\# + \mathfrak{s}l^{\natural\#})^{-1} \subset \mathfrak{p}^i \subseteq \mathfrak{s}L \\ &\Sigma_i(L) \geqq \Sigma_i(l) && \text{otherwise.} \end{aligned}$$

PROPOSITION 4.4. *Let l be well below L, of the same dimension > 2 as L, and suppose that $\mathfrak{n}l^{\#\mathfrak{q}} = \mathfrak{n}l^\#$ (cf. 4.3 for $\mathfrak{q}$). Then $l \to L$ if $Fl \cong FL$, $\mathfrak{g}l \subseteq \mathfrak{g}L$, and $\mathfrak{g}L^\# \subseteq \mathfrak{g}l^\#$.*

Proof. The proof is by induction on $N(l) = \operatorname{ord}(\mathfrak{q} \cdot \mathfrak{s}l_t) \geqq 0$, where $l = l_1 \perp \cdots \perp l_t$ is a Jordan splitting. If $N(l) = 0$, then $l \to L$ by Prop. 4.2. Suppose $N(l) > 0$; then $\dim l_t = 2$.

Now let $l = l_1 \perp \cdots \perp l_t$ be any Jordan splitting. Note that $\mathfrak{n}l_t^{\mathfrak{a}\#} = \mathfrak{n}l^\#$. We want to choose a Jordan splitting in which the last component has a particularly simple form. Let α be a norm generator of $\mathfrak{g}l^\#$ which is contained in $Q(l_t^{\mathfrak{a}\#})$, and suppose that $l_t^\# \cong \mathcal{B}_{-r}(\alpha', \beta')$. If $\alpha \sim \alpha'$, we may choose $\xi \in \mathfrak{o}$ such that $\xi^2\alpha \equiv \alpha' \bmod \mathfrak{p}\alpha'$. By Lemma 1.17, $k \cong \mathcal{A}_{-r}(\alpha' + \xi^2\alpha, \beta')$

is also the first component in some Jordan splitting of $l^{\#}$. Unless $l_t \cong \mathfrak{H}_r$, we have either $|\Delta(k)| < |\Delta(l_t^{\#})|$ or $\mathfrak{n}k \subset \mathfrak{n}l_t^{\#}$. If $\beta' \sim \alpha$, a similar thing could be done. We can now put k in canonical form: $k \cong \mathcal{B}_{-r}(\alpha'', \beta'')$, and if either α'' or $\beta'' \sim \mathfrak{n}l^{\#}$, the "reduction" process can be applied again. The reader can easily see that a sufficient number of applications leads to a Jordan splitting for $l^{\#}$ in which the first component is either a hyperbolic plane or of the form $\mathcal{B}_{-r}(\delta, \pi^{-2r}\omega\delta^{-1})$ where $\omega = 0$ or ρ and $\delta \not\sim \mathfrak{n}l^{\#}$. Let us suppose that $l_t^{\#}$ is already in one of these forms.

If $l_t^{\#}$ is hyperbolic and $\mathfrak{s}l_t^{\#} \sim \mathfrak{n}l^{\#}$, an easy computation shows that $\mathfrak{g}l^{\#} = \mathfrak{g}(l_t^{*} \perp \mathfrak{H}_{r-1})^{\#}$ whence it follows that $l \to l_t^{*} \perp \mathfrak{H}_{r-1} \to L$ by the induction hypothesis.

Now suppose $l_t^{\#} \cong \mathcal{B}_{-r}(\delta, \pi^{-2r}\omega\delta^{-1})$; this includes the case $l_t^{\#} \cong \mathfrak{H}_{-r}$, $\mathfrak{s}l_t^{\#} \not\sim \mathfrak{n}l^{\#}$. If $\delta\pi^{2r-2} \in \mathfrak{g}L$, it is easy to see that $l \to l_t^{*} \perp \mathcal{B}_{r-1}(\delta\pi^{2r-2}, \omega\delta^{-1}) \to L$ as in the previous case. Suppose $\delta\pi^{2r-2} \notin \mathfrak{g}L$. Since $\mathfrak{g}l \subseteq \mathfrak{g}L$ and $\mathfrak{n}l_t^{*} \supset \delta\mathfrak{p}^{2r}$, it is clear that $\delta\pi^{2r}$ must be a base generator of $\mathfrak{g}L$. Define

$$l' \cong l_t^{*} \perp \mathcal{A}_{r-1}(\delta\pi^{2r}, \omega\delta^{-1}\pi^{-2}).$$

The existence of a $\mathfrak{p}^{r-1}$-modular lattice on Fl_t with norm group in $\mathfrak{g}L$ (Lemma 3.2) shows that $\omega\delta^{-1}\pi^{-2} \in \mathfrak{g}L$, whence $\mathfrak{g}l' \subseteq \mathfrak{g}L$. Now let $\mathfrak{g}L = a'\mathfrak{o}^2 + b'\mathfrak{o}$. If $a' \sim \delta$, we must have $a' \in \delta\mathfrak{p}^2$ since $\mathfrak{s}L^{\#} \subset \mathfrak{s}l^{\#}$ implies that $\pi^{r-1}L^{\#} \subseteq L$ which in turn implies that $\pi^{2r-2}\mathfrak{g}L^{\#} \subseteq \mathfrak{g}L$. Clearly $\delta\mathfrak{p}^2 \subseteq \mathfrak{g}l'^{\#}$, and therefore $l \to l' \to L$ by the induction hypothesis. Finally suppose that $a' \not\sim \delta$. Choose $\alpha' \in Q(l_t^{*\#})$ with $|\alpha'| = |a'|$ and write $\alpha' = \epsilon^2 a' + d$ with $d \not\sim a'$. Since $\pi^{r-1}l_t^{*\#} \subseteq l_t^{*}$, $\pi^{2(r-1)}\alpha' \in \mathfrak{g}L$, whence $d \in \delta\mathfrak{p}^2$, and therefore $a' \in \mathfrak{g}l'^{\#}$. It is easy to show that $b' \in \delta\mathfrak{p}^2$, and so $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l'^{\#}$. Thus $l \to l' \to L$ as required.

PROPOSITION 4.5. *Let l be well below L, of the same dimension > 2 as L, and put $\mathfrak{m} = \mathfrak{n}l^{\#\mathfrak{q}}$. Suppose $\mathfrak{n}L^{\#} \subseteq \mathfrak{m}$. Then $l \to L$ if $Fl \cong FL$, $\mathfrak{g}l \subseteq \mathfrak{g}L$, and $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}(l^{\#})_{\mathfrak{m}}$.*

Proof. Let $l = l_1 \perp \cdots \perp l_t$ be a Jordan splitting; the proof is by induction on $N(l) = \operatorname{ord}(\mathfrak{q} \cdot \mathfrak{s}l_t)$. If $N(l) = 0$, the result follows by Prop. 4.2. Suppose $N(l) > 0$; then $\dim l_t = 2$. We may also suppose that $\mathfrak{m} = \mathfrak{n}l^{\#\mathfrak{q}} \subset \mathfrak{n}l^{\#}$ since otherwise $l \to L$ by Prop. 4.4. Note that $\mathfrak{g}(l^{\#})_{\mathfrak{m}} \subseteq \mathfrak{g}l^{\#}$. Let $\mathfrak{g}L = a\mathfrak{o}^2 + b\mathfrak{o}$, $\mathfrak{g}L^{\#} = a'\mathfrak{o}^2 + b'\mathfrak{o}$, and $l_t \cong \mathcal{B}_r(\alpha, \beta)$. We shall suppose, as we may, that α and β are in $a\mathfrak{o}^2$ if $\alpha \sim a$ and $|\alpha| = |\beta|$.

First Case. l_t is a hyperbolic plane. Then $l_t^{\#} \cong \mathcal{A}_{-r}(\delta, 0)$ where $\delta = a'$ if $\mathfrak{m} \sim b'$, or δ is chosen so that $\delta \sim b'$ and $\delta\mathfrak{o} \subseteq \mathfrak{p}^{-r} \cap b\mathfrak{q}^2 \subseteq \delta\mathfrak{p}^{-1}$ if $\mathfrak{m} \sim a'$. Then $l \to l' \to L$ by Prop. 4.2 where $l' \cong l_t^{*} \perp \mathcal{A}_q(\delta\pi^{2q}, 0)$ with $q = -\operatorname{ord}\mathfrak{q}$. This is easy to see if $\mathfrak{m} \sim b'$. Suppose $\mathfrak{m} \sim a'$. Choose α' in $\mathfrak{g}l^{\#\mathfrak{q}}$ with

$|\alpha'| = |a'|$. Then one can express α' in the form $\alpha' = \epsilon^2 a' + \delta'$ where $\delta' \not\sim a'$, and use the facts that $\pi^{2q}\mathfrak{g}l^{\#q} \subseteq \mathfrak{g}l$ and $\pi^{2q}\mathfrak{g}L^{\#} \subseteq \mathfrak{g}L$ to show that $|\delta'| \leqq |\delta|$. Then $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l'^{\#}$ since $\mathfrak{g}l^{\#q} \subseteq \mathfrak{g}l'^{\#}$, and the result follows.

Second Case. $\beta\pi^{-2} \notin \mathfrak{g}L$. Then β must be a base generator of $\mathfrak{g}L$ (recall the assumption above when $|\alpha| = |\beta|$, $\alpha \sim a$). By Lemma 3.2, there is a $\mathfrak{q}^{-1}$-modular lattice k on Fl_t with $\mathfrak{g}k \subseteq \mathfrak{g}L$, say $k = \mathfrak{o}x + \mathfrak{o}y \cong \mathcal{B}_q(\pi^{2q-2r}\alpha', \beta')$. By replacing x by $\pi^n x$ and y by $\pi^{-n}y$ if necessary, where $|\pi^{2n}| = |\beta'\beta^{-1}|$, we may suppose that $|\alpha'| = |\alpha|$ and $|\beta'| = |\beta|$. Now $\mathfrak{g}l_t = \mathfrak{g}\mathcal{B}_r(\alpha', \beta')$ since both norm groups are in $\mathfrak{g}L$, and β and β' are base generators of $\mathfrak{g}L$. Thus $l_t \cong \mathcal{B}_r(\alpha', \beta') \rightarrow k$, and so $l \rightarrow l'$ where $l' = l_t^{*} \perp k$, and $\mathfrak{g}l' \subseteq \mathfrak{g}L$. Next note that $\mathfrak{g}k = \pi^{2s}a\mathfrak{o}^2 + b\mathfrak{o}$ for some $s \geqq 0$. Using the procedure in 1.1 for calculating the intersection of two norm groups, one can show that $(\pi^{-2q}\mathfrak{g}L) \cap \mathfrak{g}l^{\#} = \mathfrak{g}k^{\#}$ (notice that $\mathfrak{n}k^{\#} = \mathfrak{n}l^{\#}$, so that $\pi^{2s-2q}a$ is a norm generator of $\mathfrak{g}l^{\#}$ since $k^{\#} \rightarrow l^{\#}$). Thus $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l'^{\#}$ and it follows that $l' \rightarrow L$ by Prop. 4.2.

Third Case. $\beta\pi^{-2} \in \mathfrak{g}L$, l_t not hyperbolic. Define

$$l' \cong l_t^{*} \perp \mathcal{A}_{r-1}(\alpha, \beta\pi^{-2}).$$

Then $l \rightarrow l'$, $\mathfrak{g}l' \subseteq \mathfrak{g}L$, $N(l') < N(l)$, and l' is well below L. Since $l_t^{\#}$ is not hyperbolic, $(l^{\#})_{\mathfrak{m}} \subseteq l'^{\#}$ (after an obvious identification), by domination. (Namely suppose that $v \in l^{\#}$, $Q(v) \in \mathfrak{m}$. Suppose $l_t^{\#} \cong \mathcal{B}_{-r}(\pi^{-2r}\alpha, \pi^{-2r}\beta)$ in the basis $l_t^{\#} = \mathfrak{o}x + \mathfrak{o}y$. Then $v = z + \xi x + \eta y$ where $z \in l_t^{*\#}$ and $\xi, \eta \in \mathfrak{o}$. Since $Q(z) \in \mathfrak{m}$, we must have $\xi \in \mathfrak{p}$ by Lemma 1.5). Put $\mathfrak{m}' = \mathfrak{n}l'^{\#q}$. Then $l^{\#q} \subset l'^{\#q}$ implies that $\mathfrak{m} \subseteq \mathfrak{m}'$, and so $(l^{\#})_{\mathfrak{m}} \subseteq (l'^{\#})_{\mathfrak{m}'}$. Thus $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}(l'^{\#})_{\mathfrak{m}'}$ whence $l' \rightarrow L$ as required.

LEMMA 4.6. *Suppose that* $K \cong J \perp \mathcal{B}_r(\alpha, \beta)$ *and* $K' \cong J' \perp \mathcal{B}_r(\alpha', \beta')$ *with*

$$\mathfrak{g}J = \mathfrak{g}J',\ \mathfrak{n}J \subset \mathfrak{n}K = \mathfrak{n}K' \supset \mathfrak{n}J',\ |\beta| = |\beta'|.$$

Let a be a norm generator of $\mathfrak{g}K$. *If* $\mathfrak{d}(a\alpha') \subseteq \mathfrak{d}(a\alpha)$, *then* $\mathfrak{g}K' = \mathfrak{g}K$.

Proof. Since $|\alpha| = |\alpha'| = |a|$, we may assume that $\alpha = a + \delta$, $\alpha' = a + \delta'$, with $\delta \not\sim a \not\sim \delta'$. Then $\mathfrak{d}(a\alpha') \subseteq \mathfrak{d}(a\alpha)$ means that $|\delta'| \leqq |\delta|$. Since it is clear that $\mathfrak{g}K' = \mathfrak{g}K$ if $\delta \in \mathfrak{w} = \mathfrak{p}^r + \beta\mathfrak{o}$, we may suppose that $\delta \notin \mathfrak{w}$; in particular $\delta \neq 0$.

Now let us show that we can find

$$\xi^2 a + \eta \in \mathfrak{g}J = \mathfrak{g}J',\ \text{with } \xi \in \mathfrak{p} \text{ and } |\eta| = |\delta|.$$

Since $a \in \mathfrak{g}K = \mathfrak{g}J + \mathfrak{g}\mathscr{B}_r(\alpha, \beta)$, we have $a = c + \epsilon^2(a + \delta) + b'$ where $c \in \mathfrak{g}J$, $b' \in \mathfrak{w}$, and $\epsilon \in \mathfrak{u}$ (by domination, since $\mathfrak{n}J \subset a\mathfrak{o}$). Since $\mathfrak{g}J \subseteq \mathfrak{g}K$, we may write $c = \xi^2 a + \eta$ with $\eta \not\rightharpoonup a$, and $\xi \in \mathfrak{p}$ since $\mathfrak{n}J \subset a\mathfrak{o}$. Then

$$a \equiv (\epsilon + \xi)^2 a + \epsilon^2\delta + \eta \mod \mathfrak{w}.$$

Multiply the congruence by a; a comparison of the quadratic defects on the two sides of the resulting congruence shows that $|\eta| = |\epsilon^2\delta| = |\delta|$, as required.

Next we show that $a \in \mathfrak{g}K'$. If $\delta' \in \mathfrak{w}$ $(= \mathfrak{p}^r + \beta'\mathfrak{o})$, this is obvious. Otherwise we must have $\delta' \sim \delta$, so we can find $\lambda \in \mathfrak{o}$ such that $\lambda^2\eta \equiv \delta'$ $\bmod \delta'\mathfrak{p}$ by the perfectness of the residue class field. Now

$$\lambda^2 c + a + \delta' = (1 + \lambda\xi)^2 a + (\lambda^2\eta + \delta') \in \mathfrak{g}K'$$

and so $\mathfrak{g}K'$ has a norm generator of the form $a + \delta''$ where $\delta'' \in \delta'\mathfrak{p}$; if $\mathfrak{g}K' = (a + \delta'')\mathfrak{o}^2 + b''\mathfrak{o}$, then $a + \delta' \in \mathfrak{g}K'$ implies that $|b''| \geqq |\delta'| > |\delta''|$, whence $a \in \mathfrak{g}K'$.

Now let b be a base generator of $\mathfrak{g}K$. If $b \in \mathfrak{w}$, then $b \in \mathfrak{g}K'$. If $\mathfrak{w} \subset b\mathfrak{o} \subseteq \delta\mathfrak{o}$, then $b \in \eta\mathfrak{o} \subset \mathfrak{g}K'$ since $\eta = c - \xi^2 a \in \mathfrak{g}K'$. Suppose $b \notin \mathfrak{w} + \delta\mathfrak{o}$. Since $b \in \mathfrak{g}K$,

$$b \equiv (\theta^2 a + \zeta) + \gamma^2 a \mod (\mathfrak{w} + \delta\mathfrak{o})$$

with $\theta^2 a + \zeta \in \mathfrak{g}J$, $\zeta \not\rightharpoonup a$, and $\gamma \in \mathfrak{o}$. As before one shows that $|b| = |\zeta|$, whence $b \in \zeta\mathfrak{o} \subset \mathfrak{g}K'$.

Therefore $\mathfrak{g}K = a\mathfrak{o}^2 + b\mathfrak{o} \subseteq \mathfrak{g}K'$. But $|\delta'| \leqq |\delta| \leqq |b|$ shows that $\alpha' \in \mathfrak{g}K$; clearly $\beta' \in \mathfrak{g}K$, so that $\mathfrak{g}K' \subseteq \mathfrak{g}K$. Therefore $\mathfrak{g}K' = \mathfrak{g}K$ as required.

PROPOSITION 4.7. *Let l be well below L and of the same dimension > 2 as L. Let $\mathfrak{h} = \mathfrak{g}L^\# + \mathfrak{g}l^{\#\mathfrak{q}}$, and let a' be any norm generator of $\mathfrak{g}L^\#$. Then necessary and sufficient conditions for $l \to L$ are: $Fl \cong FL$, $\mathfrak{g}l \subseteq \mathfrak{g}L$, $\mathfrak{g}L^\# \subseteq \mathfrak{g}(l^\#)_\mathfrak{h}$, and*

$$\mathfrak{g}L^\# \subseteq a'\mathfrak{o}^2 + \mathfrak{g}k^{\#\mathfrak{q}} \text{ if } \mathfrak{n}l^{\#\mathfrak{q}} \subset \mathfrak{n}L^\#, \text{ where } k^\# = (l^\#)_{\mathfrak{h}\mathfrak{o}}. \tag{10}$$

Proof. Suppose we scale l and L by a non-zero scalar α. Clearly $l \to L$ is equivalent to $l^\alpha \to L^\alpha$, and l^α is well below L^α. Also $\mathfrak{g}l^\alpha = \alpha\mathfrak{g}l$, $\mathfrak{g}L^\alpha = \alpha\mathfrak{g}L$. A straightforward computation shows that the "dual" norm groups and ideals for l^α and L^α corresponding to those in the statement of the proposition, are merely the old ones multiplied by α^{-1} (cf. the formulas preceding Lemma 1.31). Thus, for example, $\mathfrak{q}^\alpha = \alpha^{-1}\mathfrak{q}$. We may therefore suppose that l and L have been scaled so that $\mathfrak{q} = \mathfrak{o}$.

Necessity. The conditions $Fl \cong FL$ and $\mathfrak{g}l \subseteq \mathfrak{g}L$ are obvious. Since $\mathfrak{g}L^{\#} \subseteq \mathfrak{h}$, $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}(l^{\#})_{\mathfrak{h}}$ by 1.30. Therefore suppose that $\mathfrak{n}l^{\#\mathfrak{o}} \subset \mathfrak{n}L^{\#}$; this implies that $\dim l_t = 2$, if $l = l_1 \perp \cdots \perp l_t$ is a Jordan splitting.

Now $L^{\#} \to k^{\#} = P \perp R$ where $P = l_t^{*\#}$ and $R = (l_t^{\#})_{\mathfrak{h}\mathfrak{o}} \cong \mathcal{B}_{-s}(\alpha, \beta)$, say. If $\mathfrak{s}R \subseteq \mathfrak{o}$, then $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}k^{\#} = \mathfrak{g}k^{\#\mathfrak{o}}$ and (10) is proved. Therefore suppose $\mathfrak{o} \subset \mathfrak{s}R$. By Theorem 3.1, applied to the representation $k \to L$, there is a unimodular lattice $R' \cong \mathcal{B}(\alpha', \beta')$ on FR such that $L^{\#} \to P \perp R'$, $R' \to R$. If R is not hyperbolic, $\mathfrak{n}R \subseteq \mathfrak{h}\mathfrak{o} = \mathfrak{n}L^{\#}$ by 1.31; since $\mathfrak{n}P \subset \mathfrak{n}L^{\#}$, we must in fact have $\mathfrak{n}R = \mathfrak{n}L^{\#}$. If R is hyperbolic, let us choose a basis $R' = \mathfrak{o}x + \mathfrak{o}y$ with $\mathfrak{n}R' = Q(x)\mathfrak{o}$, $Q(y) = 0$. If $\mathfrak{n}R' \supset \mathfrak{n}L^{\#}$, domination shows that

$$L^{\#} \to P \perp [\mathfrak{o}(\pi x) + \mathfrak{o}y] \to P \perp [\mathfrak{o}(\pi x) + \mathfrak{o}(\pi^{-1}y)].$$

It follows that we may assume in this case too, that $\mathfrak{n}R' = \mathfrak{n}L^{\#}$. Since $\mathfrak{n}R = \mathfrak{n}R'$ if R is not hyperbolic, $|\beta'| = |\pi^{2s}\beta|$, whence $\beta' \in \mathfrak{g}k^{\#\mathfrak{o}}$. Therefore $\alpha'\mathfrak{o}^2 + \mathfrak{g}k^{\#\mathfrak{o}} = \mathfrak{g}_1$, where $\mathfrak{g}_1 = \mathfrak{g}[k^{\#\mathfrak{o}} \perp \mathcal{B}(\alpha', \beta')]$. Now α' is a norm generator of $\mathfrak{g}_2 = \mathfrak{g}[k^{\#\mathfrak{o}} \perp \mathcal{B}(\alpha', \beta')]$ whence $\mathfrak{g}_1 = \mathfrak{g}_2$ by Lemma 4.6. On the other hand, $\mathfrak{g}_2 = \mathfrak{g}(P \perp R')$ since $R^{\mathfrak{o}} = R^{\#} \to R'$, and (10) follows immediately.

Sufficiency. The proof is by induction on $N(l) = r = -\operatorname{ord}(\mathfrak{s}l^{\#})$. If $N(l) = 0$ or if $N(l) > 0$ and $\mathfrak{n}L^{\#} \subseteq \mathfrak{n}l^{\#\mathfrak{o}}$, the proposition follows by Prop. 4.5 since $\mathfrak{g}(l^{\#})_{\mathfrak{h}} \subseteq \mathfrak{g}(l^{\#})_{\mathfrak{h}\mathfrak{o}} \subseteq \mathfrak{g}l^{\#}$ and $\mathfrak{h}\mathfrak{o} = \mathfrak{n}l^{\#\mathfrak{o}}$. Therefore suppose that $\mathfrak{n}l^{\#\mathfrak{o}} \subset \mathfrak{n}L^{\#}$. Let us write $l_t \cong \mathcal{B}_r(\alpha, \beta)$ where $l = l_1 \perp \cdots \perp l_t$ is a Jordan splitting. Note that $L^{\#} \subseteq L$ since $\mathfrak{q} = \mathfrak{o}$.

1). Suppose $l_t \cong \mathcal{H}_r$, or that $|\alpha| = |\beta|$ and $\alpha \sim \alpha'$. In either case $l_t^{\#}$ is $\mathfrak{p}^{-r}$-maximal by 2.12, and so if $\delta = \Delta(l_t)$, $\mathcal{B}(\alpha', \delta\alpha'^{-1}) \to l_t^{\#}$ by 2.9. Put $l' = l_t^{*} \perp \mathcal{B}(\alpha', \delta\alpha'^{-1})$; thus $l'^{\#} \to l^{\#}$, i.e. $l \to l'$. It is easy to see that $\mathfrak{g}l'^{\#} = \alpha'\mathfrak{o}^2 + \mathfrak{g}l^{\#\mathfrak{o}}$, whence $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}l'^{\#}$ and so $l' \to L$ by Prop. 4.2. Thus $l \to L$ as required.

2). Suppose $|\alpha| = |\beta|$, $\alpha \not\sim \alpha'$. Then $\Delta(l_t) = \rho$, and 2) of 1.32 shows that $\mathfrak{g}(l_t^{\#})_{\mathfrak{h}} = b''\mathfrak{o}$ where b'' is a base generator of $\mathfrak{h}$. But then $\mathfrak{n}(l_t^{\#})_{\mathfrak{h}} \subset \mathfrak{n}L^{\#}$. a contradiction. This case is therefore impossible. We may assume from now on that $|\alpha| > |\beta|$.

3). Suppose $\beta\pi^{-2} \notin \mathfrak{g}L$. Since $|\alpha| > |\beta|$, β must be a base generator of $\mathfrak{g}L$. By Lemma 3.2, there is a unimodular lattice $l_t' \cong \mathcal{B}(\alpha', \beta')$ on Fl_t with norm group contained in $\mathfrak{g}L$. By replacing α' by $\alpha'\pi^{2n}$ and β' by $\beta'\pi^{-2n}$ for suitable $n \geqq 0$, we may suppose that $|\beta'| = |\beta|$ and therefore that

$\mathfrak{g}l_t' = (\pi^{2m}a)\mathfrak{o}^2 + \beta\mathfrak{o}$ (a is a norm generator of $\mathfrak{g}L$), and $\mathfrak{g}l_t = (\pi^{2m+2r}a)\mathfrak{o}^2 + \beta\mathfrak{o}$. It follows that $l_t \cong \mathcal{B}_r(\pi^{2r}\alpha', \beta')$ whence $l \rightarrow l'$. On the other hand,

$$\mathfrak{g}L^\# \subseteq \mathfrak{g}L \cap \mathfrak{n}l^\# = \mathfrak{g}l_t' = \mathfrak{g}l_t'^\#,$$

whence $l' \rightarrow L$ by 4.2, and finally $l \rightarrow L$.

4). Consider $\mathfrak{n}l_t^\# \supset \mathfrak{n}L^\#$. Since we may assume that $\beta\pi^{-2} \in \mathfrak{g}L$ by 3), the norm groups of $l_t' \cong \mathcal{A}_{r-1}(\alpha, \beta\pi^{-2})$ and $l' = l_t^* \perp l_t'$ are in $\mathfrak{g}L$. Now $l^{\#\mathfrak{o}} \subset l'^{\#\mathfrak{o}}$ whence $\mathfrak{h} \subseteq \mathfrak{h}' = \mathfrak{g}L^\# + \mathfrak{g}l'^{\#\mathfrak{o}}$. By 1), we may suppose that $l_t^\#$ is not hyperbolic; then by domination $(l^\#)_{\mathfrak{h}} \subseteq l'^\#$, whence $\mathfrak{g}(l^\#)_{\mathfrak{h}} \subseteq \mathfrak{g}(l'^\#)_{\mathfrak{h}'}$. By the induction hypothesis we will be finished with this case if we can show that $\mathfrak{g}L^\# \subseteq a'\mathfrak{o}^2 + \mathfrak{g}k'^{\#\mathfrak{o}}$ where $k'^\# = (l'^\#)_{\mathfrak{h}'\mathfrak{o}}$. First suppose that $\beta\pi^{-2} \in \mathfrak{h}\mathfrak{o}$. Then $\mathfrak{h}\mathfrak{o} = \mathfrak{h}'\mathfrak{o}$, so $k^\# = k'^\#$ since $(l^\#)_{\mathfrak{h}\mathfrak{o}} \subseteq l'^\# \subseteq l^\#$. Now suppose $\beta\pi^{-2} \notin \mathfrak{h}\mathfrak{o}$, so that $\mathfrak{h}'\mathfrak{o} = \beta\mathfrak{p}^{-2}$ and $\mathfrak{h}\mathfrak{o} = \alpha\mathfrak{o} = \beta\mathfrak{p}^{-1}$ (recall that $|\alpha| > |\beta|$). Domination shows that $l_t^{\#\mathfrak{o}} = (l_t^\#)_{\mathfrak{h}\mathfrak{o}} = l_t, l_t'^{\#\mathfrak{o}} = (l_t'^\#)_{\mathfrak{h}'\mathfrak{o}} = l_t'$ whence $\mathfrak{g}k^{\#\mathfrak{o}} \subseteq \mathfrak{g}k'^{\#\mathfrak{o}}$ and the result follows.

5). The last case is $\mathfrak{n}l_t^\# = \mathfrak{n}L^\#$. Then $l_t^\# \cong \mathcal{B}_{-r}(\alpha\pi^{-2r}, \beta\pi^{-2r})$ with $|\alpha\pi^{-2r}| = |a'|$. By the usual argument we can suppose that $\alpha\pi^{-2r} = a' + \delta$ with $\delta \not\rightarrow a'$. Thus $l_t^\# \cong \mathcal{B}_{-r}(a' + \delta, \beta\pi^{-2r})$; we shall assume in addition that of all such expressions for $l_t^\#$, δ has been chosen so that $\delta\mathfrak{o}$ is as small as possible; this can always be done—the only difficulty in showing the existence of such an expression occurs when $\Delta(l_t) \notin \mathfrak{o}$, in which case one uses the last part of Lemma 1.21.

Suppose that $a' + \delta \notin \mathfrak{h}$. By 1), 2) or 3) of 1.32, $\delta \in \mathfrak{p}^{-r} + \beta\mathfrak{p}^{-2r}$ since otherwise $\mathfrak{n}(l^\#)_{\mathfrak{h}} \subset \mathfrak{n}L^\#$. If $\Delta(l_t) \in \mathfrak{o}$, it is easy to see that $\delta = 0$ by the minimal character of δ. Thus $\Delta(l_t) \notin \mathfrak{o}$, and so by 1.22, we can take $\delta = \rho\beta^{-1}$ (for suitable β, of course). By formula (3) of 1.32, $\mathfrak{n}(l^\#)_{\mathfrak{h}} \subset \mathfrak{n}L^\#$. We must therefore have $a' + \delta \in \mathfrak{h}$. Since $l^{\#\mathfrak{o}} \subseteq l$ and $L^\# \subseteq L$, this implies that $a' + \delta \in \mathfrak{g}L$.

Let us define

$$K \cong l^{\#\mathfrak{o}} \perp \mathcal{B}(a', \beta), \quad K' \cong l^{\#\mathfrak{o}} \perp \mathcal{B}(a' + \delta, \beta).$$

Since $\beta \in \mathfrak{g}l^{\#\mathfrak{o}}$, it is clear that $\mathfrak{g}K = a'\mathfrak{o}^2 + \mathfrak{g}l^{\#\mathfrak{o}}$, whence $\mathfrak{g}L^\# \subseteq \mathfrak{g}K$ (since $k^\# = l^\#$), whence $\mathfrak{g}K = \mathfrak{g}L^\# + \mathfrak{g}l^{\#\mathfrak{o}} = \mathfrak{h}$. Therefore $a' + \delta$ is a norm generator of $\mathfrak{g}K$, and since $a' = (a' + \delta) + \delta$, we get $\mathfrak{g}K' = \mathfrak{g}K = \mathfrak{h}$ by Lemma 4.6. Now define $l' \cong l_t^* \perp \mathcal{B}(a' + \delta, \beta)$. Clearly $l^{\#\mathfrak{o}} \rightarrow l'^\#$, so that $\mathfrak{g}l'^\# = \mathfrak{g}K' = \mathfrak{h} \supseteq \mathfrak{g}L^\#$. By Prop. 4.2, $l' \rightarrow L$ and therefore $l \rightarrow l' \rightarrow L$ as required.

In the next theorem, the ideal $\mathfrak{q}$ is that defined in 4.3 when $\dim L > 2$;

if $\dim L = 2$, put $\mathfrak{q} = 0$. In this latter case, $l^{\#\mathfrak{q}}$ is the zero lattice, so that $\mathfrak{g}l^{\#\mathfrak{q}} = \mathfrak{n}l^{\#\mathfrak{q}} = 0$ and $\mathfrak{h} = \mathfrak{g}L^{\#}$.

THEOREM 4.8. *Let l be well below L (cf. 4.3) and of the same dimension as L. Let $\mathfrak{h} = \mathfrak{g}L^{\#} + \mathfrak{g}l^{\#\mathfrak{q}}$, let a' be a norm generator of $\mathfrak{g}L^{\#}$, and define the lattice k by $k^{\#} = (l^{\#})_{\mathfrak{h}\mathfrak{o}}$. Necessary and sufficient conditions for $l \to L$ are: $Fl \cong FL$, $\mathfrak{g}l \subseteq \mathfrak{g}L$, $\mathfrak{g}L^{\#} \subseteq \mathfrak{g}(l^{\#})_{\mathfrak{h}}$, and*

$$(11) \qquad \mathfrak{g}L^{\#} \subseteq a'\mathfrak{o}^2 + \mathfrak{g}k^{\#\mathfrak{q}} \text{ if } 0 \subset \mathfrak{n}l^{\#\mathfrak{q}} \subset \mathfrak{n}L^{\#}.$$

Proof. If $\dim L > 2$, the theorem follows by Prop. 4.7. Suppose $\dim L = 2$; since the necessity is obvious ($\mathfrak{g}L^{\#} \subseteq \mathfrak{g}(l^{\#})_{\mathfrak{h}}$ holds by 1.30), we need only consider the sufficiency. As in 4.7, we may scale l and L by the same non-zero constant. We may therefore suppose that L is unimodular. The proof proceeds inductively on $N(l) = r = \operatorname{ord}(\mathfrak{s}l)$. Since l is well below L, $\mathfrak{s}l \subseteq \mathfrak{s}L$ whence $N(l) \geqq 0$. If $N(l) = 0$, then $\mathfrak{g}l = \mathfrak{g}L$ and so $l \cong L$ by Sah's theorem.

Suppose $N(l) > 0$. The proof follows the same path as that of 4.7, when l_t in that proposition is replaced here by l. We write $l \cong \mathcal{B}_r(\alpha, \beta)$. If $l \cong \mathfrak{H}_r$, or $|\alpha| = |\beta|$, then $l^{\#}$ is maximal and so $L^{\#} \to l^{\#}$, whence $l \to L$. If $\beta\pi^{-2} \notin \mathfrak{g}L$ (and $|\alpha| > |\beta|$), it is easy to see that β is a base generator of $\mathfrak{g}L$; if $L \cong \mathcal{B}(a, b)$, then $l \cong \mathcal{B}_r(\pi^{2r}a, b)$ and $l \to L$ as in 3) of 4.7. If $\mathfrak{n}l^{\#} \supset \mathfrak{n}L^{\#}$, then $l \to \mathcal{A}_{r-1}(\alpha, \beta\pi^{-2}) \to L$ by the induction hypothesis; the proof of this is similar to 4) in 4.7, when, of course, all reference to condition (10) is omitted. Finally suppose that $\mathfrak{n}l^{\#} = \mathfrak{n}L^{\#}$. We write, as in 5) of 4.7, $l^{\#} \cong \mathcal{B}_{-r}(a' + \delta, \beta\pi^{-2r})$ and then show that $a' + \delta \in \mathfrak{h} = \mathfrak{g}L$. Clearly $l \to \mathcal{B}(a' + \delta, \beta) \cong L$. This concludes the proof of Theorem 4.8.

THEOREM 4.9. *Suppose that l is well below L and that $\nu = \dim L - \dim l > 2$. Then $l \to L$ if and only if $\mathfrak{g}l \subseteq \mathfrak{g}L$ and $\mathfrak{s}l \subseteq \mathfrak{s}L$.*

Proof. The necessity is trivial. While proving the sufficiency, we shall include the additional case $\nu = 2$, $\dim l^{*} = 2$, which is needed in Theorem 4.10. By Theorem 1.11, $Fl \to FL$. Thus if $l = l^{\ddagger} \perp l^{*}$, there is a non-defective space V such that $Fl^{\ddagger} \perp V \cong FL$ and l^{*} is in V (under suitable identification) . We next find a $\mathfrak{p}^{2r}$-modular lattice k_r on V which contains l^{*}. Suppose $l^{*} = \mathfrak{o}x$ if $\dim l^{*} = 1$, or $l^{*} = \mathfrak{o}x + \mathfrak{o}y$ if $\dim l^{*} = 2$.

If $\dim l^{*} = 0$, let K be any unimodular lattice on V and put $k_r = \mathfrak{p}^r K$.

If $\dim l^{*} = 1$, choose $x' \in V$ with $B(x, x') = 1$. Then $V = (Fx + Fx') \perp W$ for some non-defective space W. Let K be any unimodular lattice on W and put $k_r = (\mathfrak{o}x + \mathfrak{p}^{2r}x') \perp (\mathfrak{p}^r K)$.

If $\dim l^{*} = 2$, there is a splitting

$$V = (Fx + Fx') \perp (Fy + Fy') \perp W$$

by Lemma 1.13, where $B(x, x') = B(y, y') = 1$. Let K be any unimodular lattice on W (or $K = 0$ if $W = 0$). Define

$$k_r = (\mathfrak{o}x + \mathfrak{p}^{2r}x') \perp (\mathfrak{o}y + \mathfrak{p}^{2r}y') \perp (\mathfrak{p}^r K).$$

Now let $l_r = l^{\ddagger} \perp k_r$ in each case. If r is sufficiently large, $\mathfrak{g}l_r \subseteq \mathfrak{g}L$, $\mathfrak{g}L^{\#} \subseteq \mathfrak{p}^{-2r} \subseteq \mathfrak{g}l_r^{\#}$, and l_r and L will satisfy (5) and (6) of Prop. 4.2 (see (9) in 4.3). Thus $l \to l_r \to L$ as required.

THEOREM 4.10. *Suppose that L is non-defective and that l is a lattice in FL with* $\dim L - \dim l = 2$.

(i) *Suppose l is non-defective. Let* $\mathfrak{p}^r = (\mathfrak{s}l^{\#} + \mathfrak{s}L^{\#})^{-1}$ *and choose a non-zero α in* $\mathfrak{p}^{2r}(\mathfrak{n}L)^{-1}$ *such that* $Fl \perp \langle \alpha \rangle \to FL$. *Put* $\delta = \Delta(l) + \Delta(L)$ *and choose s large enough so that* $\mathfrak{p}^s + \delta\alpha^{-1}\mathfrak{p}^{2s} \subseteq \alpha\mathfrak{o}$. *Then $l \to L$ if and only if $l \perp P \to L$ where* $P \cong \mathcal{B}_s(\alpha, \delta\alpha^{-1}\pi^{2s})$.

(ii) *Suppose that l is defective, or that* $\mathfrak{n}L^{\#} \subseteq \mathfrak{n}l^{\#}$. *If l is well below L, then $l \to L$ if and only if* $\mathfrak{s}l \subseteq \mathfrak{s}L$ *and* $\mathfrak{g}l \subseteq \mathfrak{g}L$.

Remarks. There is no loss of generality in the assumption that l is in FL since $Fl \to FL$ is a necessary condition for $l \to L$.

This theorem solves completely the case $\nu = \dim L - \dim l = 2$ and l well below L. For (i) reduces $\nu = 2$ to $\nu = 0$ for which there is a solution when $l \perp P$ is well below L, and this follows from the fact that $\mathfrak{p}^s = \mathfrak{s}P \subseteq \mathfrak{p}^r$ since $\mathfrak{p}^r \subseteq \mathfrak{s}L$ and therefore

$$\mathfrak{p}^s \subseteq \alpha\mathfrak{o} \subseteq \mathfrak{p}^{2r}(\mathfrak{n}L)^{-1} \subseteq \mathfrak{p}^{2r}(\mathfrak{s}L)^{-1} \subseteq \mathfrak{p}^r.$$

Proof of (i). The sufficiency is trivial. Suppose that $l \to L$. By Theorem 3.1, there is a binary $\mathfrak{p}^r$-modular lattice k such that $l \perp k \to L$. Since $\mathfrak{n}k \subseteq \mathfrak{n}L$, $\mathfrak{p}^{2r}(\mathfrak{n}L)^{-1} \subseteq (\mathfrak{n}k^{\#})^{-1}$, and so $P \to k$ by 2.11 and 2.9. Therefore $l \perp P \to l \perp k \to L$ as required.

Proof of (ii). This time the necessity is trivial. Suppose $\mathfrak{s}l \subseteq \mathfrak{s}L$ and $\mathfrak{g}l \subseteq \mathfrak{g}L$. The case when l is defective was done in the proof of Theorem 4.9. Suppose that $l^{*} = 0$ and that $\mathfrak{n}L^{\#} \subseteq \mathfrak{n}l^{\#}$. Let M be a $(\mathfrak{n}l^{\#})^{-1}$-maximal lattice such that $Fl \perp FM = FL$. We shall show that $l \perp M \to L$.

Note that $\mathfrak{s}M \subseteq (\mathfrak{n}l^{\#})^{-1}$. Thus since $\mathfrak{s}l^{\#} \subseteq \mathfrak{n}l^{\#}$, M is the last component (or at least part of it) in some Jordan splitting of $l \perp M$. Also $\mathfrak{s}L^{\#} \subseteq \mathfrak{n}l^{\#}$,

and it follows that $l \perp M$ is well below L. Let $\mathfrak{q} = \mathfrak{s}l^{\#}$ and $\mathfrak{m} = \mathfrak{n}(l \perp M)^{\#\mathfrak{q}}$. By Prop. 4.5, it will suffice to show that $\mathfrak{g}(l \perp M) \subseteq \mathfrak{g}L$ and

$$\mathfrak{g}L^{\#} \subseteq \mathfrak{g}((l \perp M)^{\#})_{\mathfrak{m}} = \mathfrak{g}l^{\#} + \mathfrak{g}(M^{\#})_{\mathfrak{m}}.$$

Clearly $\mathfrak{g}(l \perp M) \subseteq \mathfrak{g}L$ since $\mathfrak{g}M \subseteq (\mathfrak{n}l^{\#})^{-1} \subseteq (\mathfrak{s}L^{\#})^{-1}$. We shall finish the proof by showing that $\mathfrak{n}l^{\#} \subseteq \mathfrak{g}l^{\#} + \mathfrak{g}(M^{\#})_{\mathfrak{m}}$.

Now $M^{\#}$ is $\mathfrak{a}$-maximal where $\mathfrak{a} = (\mathfrak{s}M)^{-2}(\mathfrak{n}l^{\#})^{-1} \supseteq \mathfrak{n}l^{\#}$. Thus $M^{\#}$ contains a $\mathfrak{n}l^{\#}$-maximal lattice M', and $M' \subseteq (M^{\#})_{\mathfrak{m}}$ since $\mathfrak{n}l^{\#} \subseteq \mathfrak{m}$. Since M' is maximal, $\mathfrak{g}M' = \mathfrak{n}M'$ (we leave this to the reader to check). Therefore $\mathfrak{p}\mathfrak{n}l^{\#} \subseteq \mathfrak{g}M' \subseteq \mathfrak{g}(M^{\#})_{\mathfrak{m}}$ and so $\mathfrak{n}l^{\#} \subseteq \mathfrak{g}l^{\#} + \mathfrak{g}(M^{\#})_{\mathfrak{m}}$ as required.

THEOREM 4.11. *Let L be a non-defective lattice and l a lattice with $\dim L - \dim l = 1$. Suppose $l = l^{\dagger} \perp l^{*}$ and $\delta = \Delta(l^{\dagger}) + \Delta(L)$. Let r be an integer such that*

(12) $$\mathfrak{p}^{r} \subset (\mathfrak{n}l^{*}) \cap (\mathfrak{n}l^{\dagger\#} + \mathfrak{n}L^{\#})^{-1}.$$

Then $l \to L$ if and only if

$$l^{\dagger} \perp \mathfrak{A}_{r}(\alpha, \delta\alpha^{-1}\pi^{2r}) \to L$$

where $l^{*} \cong \langle\alpha\rangle$.

Remark. Let $P \cong \mathfrak{A}_{r}(\alpha, \delta\alpha^{-1}\pi^{2r})$. By (12),

$$\mathfrak{s}P = \mathfrak{p}^{r} \subset (\mathfrak{n}l^{\dagger\#} + \mathfrak{n}L^{\#})^{-1} \subseteq (\mathfrak{s}l^{\dagger\#} + \mathfrak{s}L^{\#})^{-1}$$

so that $\mathfrak{s}P$ is contained in both $(\mathfrak{s}L^{\#})^{-1}$ and $(\mathfrak{s}l^{\dagger\#})^{-1}$ (if $l^{\dagger} \neq 0$). Thus it is clear that $l^{\dagger} \perp P$ is well below L if l (or $l^{\dagger}$) is, and in this case Theorem 4.8 can be used to see whether or not $l^{\dagger} \perp P \to L$.

Proof. The sufficiency is trivial, so suppose $l \to L$. By Theorem 3.1, there is a binary $\mathfrak{p}^{r}$-modular lattice R such that $l^{*} \to R$ and $l^{\dagger} \perp R \to L$. Choose a basis $R = \mathfrak{o}x + \mathfrak{o}y$ such that $B(x, y) = \pi^{r}$ and $Q(x)Q(y)(B(x,y))^{-2} = \delta$. Now $l^{*} \to R$ is equivalent to $\alpha \in Q(R)$ and so $\alpha = Q(\xi x + \eta y)$ for some $\xi, \eta \in \mathfrak{o}$. By domination (if $\delta = 0$, note that $\alpha \notin \mathfrak{p}^{r}$ by (12)),

$$|\alpha| = \sup\{|Q(\xi x)|, |Q(\eta y)|\}.$$

By interchanging x and y if necessary, we may assume that $|\alpha| = |Q(\xi x)|$. Now $\mathfrak{p}^{2r} \subset \mathfrak{n}l^{*} \cdot (\mathfrak{n}l^{\dagger\#} + \mathfrak{n}L^{\#})^{-1}$ and so

(13) $$\mathfrak{n}l^{\dagger\#} + \mathfrak{n}L^{\#} \subset \mathfrak{p}^{-2r}\mathfrak{n}l^{*} \subseteq Q(x)\mathfrak{p}^{-2r}.$$

Now suppose that $\xi\mathfrak{o} = \mathfrak{p}^{s}$. Then $l^{*} \to S = \mathfrak{o}(\pi^{s}x) + \mathfrak{o}(\pi^{-s}y)$; also domina-

tion and (13) show that $L \to l^{\natural\#} \perp S$. Replacing R by S, we may suppose that ξ is a unit, whence $|\alpha| = |Q(x)|$.

If $|Q(y)| \leqq |Q(x)|$, then $P \cong R$ by Lemma 1.20, and so $l^\natural \perp P \to L$. Suppose $|Q(y)| > |Q(x)|$; then $Q(y) \not\to Q(x)$ and $\delta \notin \mathfrak{o}$ (since $Q(x) \notin \mathfrak{s}P$). Let t be the smallest integer such that $|\pi^{-2t}Q(x)| > |Q(y)|$, and put $T = \mathfrak{o}(\pi^{-t}x) + \mathfrak{o}y$. Domination shows that T is $Q(x)\mathfrak{p}^{-2t}$-maximal. Also

$$L^\# \to l^{\natural\#} \perp (\mathfrak{o}(\pi^{-r}x) + \mathfrak{o}(\pi^{-r+t}y)) = l^{\natural\#} \perp T^\#$$

by (13), domination, and the choice of t. Now $\mathfrak{n}P = Q(y)\mathfrak{o} \subset \mathfrak{n}T$, so that $P \to T$ by 2.9. Therefore $l^\natural \perp P \to l^\natural \perp T \to L$ as required.

UNIVERSITY OF NOTRE DAME.

REFERENCES.

[1] C. Arf, "Untersuchungen über quadratische Formen in Körpern der Charakteristik 2," (Teil. I.), *Journal für die Reine und Angewandte Mathematik*, vol. 183 (1940), pp. 148-167.

[2] ———, "Über arithmetische Aequivalenz quadratische Formen in Potenzreihenkörpern über einem vollkommen Körpern der Charakteristik 2," *Revue de la Faculté des Sciences de l'Université d'Istanbul*, vol. 8 (1943), pp. 297-327.

[3] M. Eichler, *Quadratische Formen und orthogonale Gruppen*, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1952.

[4] O. T. O'Meara, "Integral equivalence of quadratic forms in ramified local fields," *American Journal of Mathematics*, vol. 79 (1957), pp. 157-186.

[5] ———, *Introduction to Quadratic Forms*, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1963.

[6] C. R. Riehm, "On the integral representations of quadratic forms over local fields," *American Journal of Mathematics*, vol. 86 (1964), pp. 25-62.

[7] C.-H. Sah, *Integral equivalence of quadratic forms in local fields of characteristic 2*, Princeton University thesis, 1959.

[8] ———, "Quadratic forms over fields of characteristic 2," *American Journal of Mathematics*, vol. 82 (1960), pp. 812-830.

[9] E. Witt, "Theorie der quadratische Formen in beliebigen Körpern," *Journal für die Reine und Angewandte Mathematik*, vol. 176 (1937), pp. 31-44.

ON STARLIKE FUNCTIONS OF ORDER α.

By ALBERT SCHILD.

1. **Introduction.** Let $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ be analytic for $|z| < 1$. We say that $f(z)$ is starlike of order α in the unit circle, if $Re\{zf'(z)/f(z)\} > \alpha$ for all $|z| < 1$, for a given α, $0 \leqq \alpha \leqq 1$, and we denote by S_α the class of all such functions, for a given α. The class S_0 has been studied extensively and many of the results are classical. Attention to the importance of class $S_{\frac{1}{2}}$ was first drawn by A. Marx [3] and E. Strohäcker [7], who showed that if $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ maps the unit circle onto a convex region, then $f(z) \in S_{\frac{1}{2}}$. Somewhat later, R. F. Gabriel [1] showed that functions of class $S_{\frac{1}{2}}$ played an important role in the solution of certain differential equations. Some results, both of a geometric and analytic nature, about functions of class $S_{\frac{1}{2}}$ were obtained by the author [6] and T. H. MacGregor [2]. The general class S_α was first introduced by M. S. Robertson [5]. Recently, some results on functions of class S_α were obtained by E. P. Merkes et al [4]. In the present paper, some further results about functions of class S_α will be obtained. It is interesting to note that the class $S_{\frac{1}{2}}$ plays a rather unique role, since in many results in this paper α appears in the form $(2\alpha - 1)$.

2. **A coefficient theorem.** Let $f(z) = z + \sum_{n=2}^{\infty} a_n z^n \in S_\alpha$. Then

$$|a_n| \leqq \frac{\prod_{k=2}^{n} (k - 2\alpha)}{(n-1)!}.$$

Proof. Let $p(z) = \dfrac{\{zf'(z)/f(z)\} - \alpha}{1-\alpha} = 1 + c_1 z + c_2 z^2 + \cdots$. Since $p(z)$ is regular and $Re\{p(z)\} > 0$ for $|z| < 1$, therefore, by a well-known lemma, we have $|c_n| \leqq 2$ for $n = 1, 2, 3, \cdots$. Comparing coefficients, we have:

$$(n-1)a_n = (1-\alpha)[c_{n-1} + a_2 c_{n-2} + \cdots + a_{n-1} c_1] \text{ for } n = 2, 3, 4, \cdots.$$

Therefore, using the lemma, we get:

$$(n-1)|a_n| \leqq (2-2\alpha)[1 + |a_2| + |a_3| + \cdots + |a_{n-1}|] \quad n = 2, 3, 4, \cdots.$$

Received March 13, 1964.

In particular:

$$
\begin{aligned}
|a_2| &\leqq (2-2\alpha) \\
2\,|a_3| &\leqq (2-2\alpha)[1+|a_2|] \\
&\leqq (2-2\alpha)[1+(2-2\alpha)] \\
&\leqq (2-2\alpha)(3-2\alpha) \\
\text{i.e. } |a_3| &\leqq (2-2\alpha)(3-2\alpha)/2!
\end{aligned}
$$

Simple induction shows that $|a_n| \leqq \dfrac{\Pi\,(k-2\alpha)}{(n-1)!}$ $n=2,3,4,\cdots$. These estimates for $|a_n|$ are sharp. They are attained by the functions

$$f(z)=z(1-z)^{-(2-2\alpha)} \in S_\alpha.$$

Note that when $\alpha=\frac{1}{2}$, $|a_n|\leqq 1$, a result which was obtained by the author [6], and $|a_n|<1$ when $\alpha>\frac{1}{2}$. Notice also that when $\alpha=0$, we get the well-known result $|a_n|\leqq n$.

3. Some distortion theorems.

THEOREM 3.1. For all $f(z)\in S_\alpha$, the domain of values of $\{zf'(z)/f(z)\}$ is the circle with the line segment from $\dfrac{1+(2\alpha-1)|z|}{1+|z|}$ to $\dfrac{1-(2\alpha-1)|z|}{1-|z|}$ as a diameter.

Proof. Let $G(z)=\{zf'(z)/f(z)\}-\alpha$ and

$$
\begin{aligned}
H(z) &= [G(z)-(1-\alpha)]/[G(z)+(1-\alpha)] \\
&= [\{zf'(z)/f(z)\}-1]/[\{zf'(z)/f(z)\}-(2\alpha-1)]
\end{aligned}
$$

$H(z)$ is regular for $|z|<1$, $H(0)=0$ and $|H(z)|<1$ for $|z|<1$. Therefore, by the Lemma of Schwartz, we have

(1) $|[\{zf'(z)/f(z)\}-1]/[\{zf'(z)/f(z)\}-(2\alpha-1)]|<|z|$ for $|z|<1$.

But the domain defined by this inequality is the circle of Apollonius with the line segment from $\dfrac{1+(2\alpha-1)|z|}{1+|z|}$ to $\dfrac{1-(2\alpha-1)|z|}{1-|z|}$ as a diameter.

Inequality (1) can also be expressed in the following form:

(2) $[\{zf'(z)/f(z)\}-1]/[\{zf'(z)/f(z)\}-(2\alpha-1)]=-z\phi(z)$,

where $\phi(z)$ is analytic and $|\phi(z)|\leqq 1$ for $|z|<1$. Solving for $\{zf'(z)/f(z)\}$ we obtain:

$$
\begin{aligned}
\{zf'(z)/f(z)\} &= [1+(2\alpha-1)z\phi(z)]/[1+z\phi(z)] \\
&= (2\alpha-1)+\frac{2(1-\alpha)}{1+z\phi(z)}.
\end{aligned}
\tag{3}
$$

Note also that some kind of a converse is true, i.e. if (3) is satisfied then $Re\{zf'(z)/f(z)\} > 2\alpha - 1) + 2(1-\alpha)/2 = \alpha$, i.e. $f(z) \in S_\alpha$.

THEOREM 3.2. *$f(z) \in S_\alpha$ if and only if*

$$f(z) = z \cdot \exp\left\{ -2(1-\alpha) \int_0^z \frac{\phi(t)\,dt}{1+t\phi(t)} \right\}.$$

Proof. From Theorem 3.1 we have: $\dfrac{zf'(z)}{f(z)} = (2\alpha - 1) + \dfrac{2(1-\alpha)}{1+z\phi(z)}$ or:

$$\frac{f'(z)}{f(z)} = \frac{2\alpha-1}{z} + \frac{2(1-a)}{z(1+z\phi(z))} = \frac{1}{z} - \frac{2(1-\alpha)\phi(z)}{1+z\phi(z)}.$$

Integration gives the required result. Conversely, suppose that $f(z)$ has such a representation. That $f(z) \in S_\alpha$ follows from differentiation, simple manipulation, and Theorem 3.1.

4. The radius of convexity of functions of class S_α. We define r_α, the radius of convexity of functions of class S_α, as the upper bound of the radii, r, of circles $|z| < r$, which are mapped by any function $f(z) \in S_\alpha$ onto a convex region. It is well-known that if $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ is analytic and univalent for $|z| < 1$, then the condition: $Re\{\dfrac{zf''(z)}{f'(z)} + 1\} > 0$ for $|z| < r$, is necessary and sufficient for $f(z)$ to map $|z| < r$ onto a convex region. It is also well-known that if $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$ is analytic and univalent for $|z| < 1$, then $r = 2 - \sqrt{3} = .268 \cdots$. It is interesting to note that this is an exact estimate, attained by the function

$$f(z) = \frac{z}{(1-z)^2} = z + \sum_{n=2}^{\infty} n z^n \in S_0,$$

and hence $r_0 = 2 - \sqrt{3}$.

MacGregor [2] has shown that the radius of convexity for functions of class $S_{\frac{1}{2}}$ is given by $r_{\frac{1}{2}} = (2\sqrt{3} - 3)^{\frac{1}{2}} = .68 \cdots$. The estimate is exact and is attained by the function $f(z) = z(1 - 2bz + z^2)^{-\frac{1}{2}} \in S_{\frac{1}{2}}$, where $b = (2 - \sqrt{3})^{\frac{1}{2}}$. The radius of convexity, r_α, for the class S_α will now be investigated.

Differentiation and simple manipulation of (3) yields:

$$\text{(4)} \qquad \frac{zf''(z)}{f'(z)} + 1 = \frac{zf'(z)}{f(z)} - 2(1-\alpha)\,\frac{(\phi(z) + z\phi'(z))}{(1+z\phi(z))^2}\,\frac{f(z)}{f'(z)}.$$

Substitution from (3) into (4) and simplification yields:

$$\frac{zf''(z)}{f'(z)}+1=\frac{1+(6\alpha-4)z\phi(z)+(2\alpha-1)^2z^2\phi^2(z)-2(1-\alpha)z^2\phi'(z)}{(1+z\phi(z))[1+(2\alpha-1)z\phi(z)]} \tag{5}$$

The function $\phi(z)=\dfrac{z-b}{1-bz}$, $-1<b<1$, maps $|z|<1$ onto itself and $\phi(z)\equiv 1$ when $b=-1$, and $\phi(z)\equiv -1$ if $b=1$. Hence $\phi(z)=\dfrac{z-b}{1-bz}$, $-1\leqq b\leqq 1$, satisfies all the conditions required in Theorem 3.1, equation (2).

Substituting into (5) we get after simplification:

$$\frac{zf''(z)}{f'(z)}+1= \tag{6}$$

$$\frac{(2\alpha-1)^2z^4+2b(1+\alpha-4\alpha^2)z^3+2(-3+2\alpha+2\alpha^2b^2)z^2+2b(1-3\alpha)z+1}{(1-2bz+z^2)[1-2\alpha bz+(2\alpha-1)z^2]}$$

if $-1<b<1$, and

$$\frac{zf''(z)}{f'(z)}+1=\frac{(2\alpha-1)^2z^2+(6\alpha-4)z+1}{(1+z)[1+(2\alpha-1)z]},\ \text{if } b=-1, \phi(z)\equiv 1. \tag{7}$$

We consider now the least positive $r=r(\alpha,b)$, such that for some z, $|z|=r$, we have $\dfrac{zf''(z)}{f'(z)}+1=0$, i.e. for this z we have from (6),

$$\begin{aligned}(2\alpha-1)^2z^4+2b(1+\alpha-4\alpha^2)z^3+2(-3+2\alpha+2\alpha^2b^2)z^2\\ +2b(1-3\alpha)z+1=0\end{aligned} \tag{8}$$

if $|b|<1$, and from (7),

$$(2\alpha-1)^2z^2+(6\alpha-4)z+1=0 \text{ if } b=-1. \tag{9}$$

The above equations can now be used to find upper bounds for r_α, since obviously $r_\alpha\leqq r(\alpha,b)$.

Equation (8) can be considered to define z implicitly as a function of b, for a given value of α. Taking the partial derivative of equation (8) with respect to b and solving for b, we get:

$$b=\frac{(4\alpha^2-\alpha-1)z^2+(3\alpha-1)}{4\alpha^2z}, \tag{10}$$

provided that the absolute value of the R. H. S. of (10) is less than 1, otherwise $b=\pm 1$. For the value of b, determined by (10), after substitution into (8) and simplification, we get

$$(8\alpha^2-3\alpha-1)z^4-(8\alpha^2-2\alpha+2)z^2+(5\alpha-1)=0 \tag{11}$$

The four roots of this equation are given by:

$$z = \pm \left\{ \frac{(4\alpha^2 - \alpha + 1) \pm 4\alpha\sqrt{\alpha^2 - 3 + 2}}{8\alpha^2 - 3\alpha - 1} \right\}^{\frac{1}{2}} \tag{12}$$

Considering—for a certain range of α, to be determined later—

$$\begin{aligned} r(a, b) = z &= \left\{ \frac{(4\alpha^2 - \alpha + 1) - 4\alpha\sqrt{\alpha^2 - 3 + 2}}{8\alpha^2 - 3\alpha - 1} \right\}^{\frac{1}{2}} \\ &= \left\{ \frac{5\alpha - 1}{(4\alpha^2 - \alpha + 1) + 4\alpha\sqrt{\alpha^2 - 3 + 2}} \right\}^{\frac{1}{2}}, \end{aligned} \tag{13}$$

we find that for this value of z we have from (10):

$$b = \frac{1}{\alpha(5\alpha - 1)} \{4\alpha(2\alpha - 1) + (3\alpha - 1)\sqrt{\alpha^2 - 3\alpha + 2}\} \cdot r(\alpha, b), \tag{14}$$

provided that for these values of α and $r(\alpha, b)$, $|b| < 1$. Notice that if b is real, z is real.

If for a given α and the corresponding value of $r(\alpha, b)$ as determined by (13), the value of b, as determined by (14) would be such that $b < -1$, then equation (9) would be used to calculate the value of $r(\alpha, -1)$. From equation (19) we get:

$$\begin{aligned} z = r(\alpha, -1) &= \frac{(2 - 3\alpha) - \sqrt{5\alpha^2 - 8\alpha + 3}}{(2\alpha - 1)^2} \\ &= \frac{1}{(2 - 3\alpha) - \sqrt{5\alpha^2 - 8\alpha + 3}}. \end{aligned} \tag{15}$$

An examination of equations (13) and (14) shows that b decreases monotonically with $\alpha, b = 1$ when $\alpha = 1$. To determine for what values of α, equation (13) should be used and for what values equation (15), we equate the values of z in both equations. Some manipulation and simplification shows that α satisfies:

$$20\alpha^4 - 52\alpha^3 + 15\alpha^2 + 12\alpha - 4 = 0. \tag{16}$$

The smallest positive α satisfying equation (16) is: $\alpha = \alpha_0 = .335\ldots$, i.e. for $\alpha > \alpha_0$, we use equation (13) to find $r(\alpha, b)$ and for $\alpha \leqq \alpha_0$, we use equation (15) to find $r(\alpha, -1)$.

To determine how good these values of $r(\alpha, b)$ are as estimates for r_α, we observe:

for $\alpha=0$, equation (15) yields $r(0,-1)=2-\sqrt{3}=r_0$

for $\alpha=\frac{1}{2}$, equation (13) yields $r(\frac{1}{2},\sqrt{2-\sqrt{3}})=(2\sqrt{3}-3)^{\frac{1}{2}}=r_{\frac{1}{2}}$

for $\alpha=1$, equation (13) yields $r(1,1)=1=r_1$.

For these three values of α, we note that $r_\alpha=r(\alpha,b)$. It seems likely that $r_\alpha=r(\alpha,b)$—where $r(\alpha,b)$ is given by equations (13) and (15) respectively—is generally true. In that case, a set of extremal functions would be: $f(z)=z(1-2bz+z^2)^{-(1-\alpha)}$, where b is given by (14), if $\alpha>.335\cdots$, and $f(z)=z(1+z)^{-2(1-\alpha)}$, if $\alpha\leqq.335\cdots$.

TEMPLE UNIVERSITY.

REFERENCES.

[1] R. F. Gabriel, "The Schwarzian derivative and convex functions," *Proceedings of the American Mathematical Society*, vol. 6 (1955), pp. 58-66.

[2] T. H. MacGregor, "The radius of convexity of starlike functions or order ½," *Proceedings of the American Mathematical Society*, vol. 14 (1963), pp. 71-76.

[3] A. Marx, "Untersuchungen Über Schlichte Abbildungen," *Mathematische Annalen*, vol. 107 (1932-33), pp. 40-67.

[4] E. P. Merkes, M. S. Robertson and W. T. Scott, "On products of starlike functions," *Proceedings of the American Mathematical Society*, vol. 13 (1962), pp. 960-964.

[5] M. S. Robertson, "On the theory of univalent functions," *Annals of Mathematics*, vol. 37 (1936), pp. 374-408.

[6] A. Schild, "On a class of univalent, starshaped mappings," *Proceedings of the American Mathematical Society*, vol. 9 (1958), pp. 751-757.

[7] E. Strohäcker, "Beiträge Zur Theorie Der Schlichten Funktionen," *Mathematische Zeitschrift*, vol. 37 (1933), pp. 356-380.

THE ČECH COHOMOLOGY OF PARACOMPACT PRODUCT SPACES.

By Ronald C. O'Neill.*

1. Introduction. Our main purpose in this paper is to prove the following two theorems:

THEOREM 1. *If $X \times Y$ is a paracompact Hausdorff space and G is an abelian group, then for each integer $n \geqq 0$,*

$$\check{H}^n(X \times Y; G) \cong \sum_{q=0}^{n} \check{H}^q(X; \check{H}^{n-q}(Y; G)).$$

THEOREM 2. *Let $X \times Y$ be a paracompact Hausdorff space and let L be a principal ideal domain. If X is compact or if $\check{H}^q(Y; L)$ is finitely generated over L in each dimension q, then $\check{H}^q(X \times Y; L)$ is given by the Künneth rule: there is an exact sequence*

$$0 \to \sum_{q=0}^{n} \check{H}^q(X; L) \otimes_L \check{H}^{n-q}(Y; L) \to \check{H}^n(X \times Y; L) \to \sum_{q=0}^{n} \check{H}^{q+1}(X; L) *_L \check{H}^{n-q}(Y; L) \to 0.$$

Theorem 2 answers in the affirmative a question by A. Borel. In [1, pp. 243-244], Borel gives a proof of Theorem 2 with the additional hypotheses: X is compact, clc_L, of finite cohomological dimension over L and satisfies the first axiom of countability. He asks whether finite dimensionality and first countability can be dispensed with. Further, Borel comments that examples show the necessity of an assumption of the type "X is compact clc_L." We remark that such a condition insures that $\check{H}^q(X; L)$ is a finitely generated L-module for each integer $q \geqq 0$ [14].

The proofs we give here rely on homotopical methods, generalized to k-homotopy classes of k-maps. In this we depend heavily on the theory of k-maps as developed by R. Brown [2] and on a theorem due to P. J. Huber [5]. In § 2 we list several of Brown's results on k-maps and some of their elementary consequences for k-homotopy classes of k-maps. (See § 2 for

Received March 16, 1964.

* The author gratefully acknowledges support by NSF Grant GP-1815 during the preparation of this paper.

definitions.) The theorem of Huber and some results due to R. Thom and to J. C. Moore are also listed in § 2. The proof of Theorem 1 is given in § 3, and the proof of Theorem 2 is given in § 4. It is perhaps an interesting feature of our proofs that we make no use of spectral sequences.

The author takes this opportunity to thank Frank Raymond for pointing out the problem treated here and for several helpful conversations on the subject of this paper.

2. Preliminary considerations. All topological spaces considered in this paper will be Hausdorff spaces with basepoints; all maps are to respect basepoints. Let X and Y be spaces and let $f\colon X \to Y$ be a basepoint-preserving function. Following R. Brown [2] we say that f is *k-continuous*, or that f is a *k-map*, if the restriction of f to each compact subset of X is a continuous function. Clearly, the composition of two k-maps is again a k-map. A k-map is a *k-homeomorphism* if it is one-to-one, onto and has a k-continuous inverse.

Let W and X be spaces and let $w_0 \in W$ and $x_0 \in X$ be the basepoints. Then $W \times X$ denotes the usual cartesian product space with basepoint (w_0, x_0). The space $W \divideontimes X$ is the identification space obtained from $W \times X$ by collapsing the subspace $(W \times x_0) \cup (w_0 \times X)$ to a single point, which we take as the basepoint of $W \divideontimes X$. We denote by $w \divideontimes x$ the point of $W \divideontimes X$ determined by $(w, x) \in W \times X$. If $f\colon W \to Y$ and $g\colon X \to Z$ are k-maps then the function $f \times g\colon W \times X \to Y \times Z$ defined by

$$(f \times g)(w, x) = (f(w), g(x)),\ w \in W, x \in X,$$

is a k-map. Likewise, the function $f \divideontimes g\colon W \divideontimes X \to Y \divideontimes Z$ induced by $f \times g$ is a k-map.

Let I denote the closed unit interval $[0, 1]$ with 1 as basepoint and let X and Y be spaces with basepoints x_0, y_0, respectively. If $f, g\colon X \to Y$ are k-maps, then $F\colon f \sim g$ is a *k-homotopy* from f to g if $F\colon X \times I \to Y$ is a k-map such that

$$F(x, 0) = f(x),\ F(x, 1) = g(x),\ x \in X,$$

and

$$F(x_0, t) = y_0,\ t \in I.$$

The relation of k-homotopy is an equivalence relation on the set of k-maps $f\colon X \to Y$. If there is a k-map $h\colon Y \to X$ such that $hf\colon X \to X$ and $fh\colon Y \to Y$ are homotopic to identity maps then f is said to be a *k-homotopy equivalence*, X and Y are said to be of the same *k-homotopy type*, and we

write $X \sim Y$. We write $X \simeq Y$ to indicate that X and Y have the same homotopy type in the usual sense.

The set of all (basepoint-preserving) k-maps of X into Y with the compact-open topology is denoted Y^X; the basepoint of Y^X is the constant k-map of X into the basepoint of Y. When X is compact or, more generally, whenever X is a k-space [6, p. 230], all k-maps $f\colon X \to Y$ are continuous and the symbol Y^X has its more usual meaning. For example, Y^{S^q}, the space of k-maps of the euclidean q-sphere into Y, and $\Omega^q Y$, the space of loops in Y at the basepoint, are the familiar function spaces. In this connection we recall:

(2.1) *Y^{S^q} and $\Omega^q Y$ are homeomorphic for all spaces Y and all $q \geqq 0$.*

Let W, X, Y, Z be arbitrary spaces. We list here several important properties of k-maps and spaces of k-maps [2], (cf. also [12]):

(2.2) *The exponential map $e\colon Y^{W \divideontimes X} \to (Y^X)^W$ defined by*

$$e(f)(w)(x) = f(w \divideontimes x),\ w \in W,\ x \in X,$$

is a homeomorphism. Consequently $(Y^W)^X$ and $(Y^X)^W$ are naturally homeomorphic spaces.

(2.3) *If $f\colon W \to X$ and $g\colon Y \to Z$ are k-maps, they induce functions $Y^f\colon Y^X \to Y^W$ and $g^X\colon Y^X \to Z^X$, where*

$$Y^f(h) = hf, \quad G^X(h) = gh, \quad h \in Y^X.$$

Y^f is a continuous function and g^X is k-continuous.

(2.4) *The natural map $r\colon (Y \times Z)^X \to Y^X \times Z^X$ is a homeomorphism.*

(2.5) *Every k-homotopy $f \sim u\colon W \to X$ induces a k-homotopy $Y^f \sim Y^u\colon Y^X \to Y^W$ and every k-homotopy $g \sim v\colon Y \to Z$ induces a k-homotopy $g^X \sim v^X\colon Y^X \to Z^X$. Thus if $W \sim X$, then $Y^X \sim Y^W$, and if $Y \sim Z$, then $Y^X \sim Z^X$.*

The set of (basepoint-preserving) k-homotopy classes of k-maps $h\colon X \to Y$ is denoted $[X;Y]$; the k-homotopy class of $h\colon X \to Y$ is denoted $[h]$. If X is any (nonempty) space we write X^+ to represent the disjoint union of X with a one-point space X_0; the basepoint of X^+ is X_0. Then $[X^+;Y]$ is clearly in natural one-to-one correspondence with the set of k-homotopy classes of those k-maps of X into Y that do not respect basepoints. Note also that if

P is a space consisting of a single point, then $[X;Y]$ is in natural one-to-one correspondence with $[P^{\dagger};Y^X]$. This observation together with (2.2) yields:

(2.6) *$[W;Y^X]$, $[W \circledast X;Y]$ and $[X;Y^W]$ are in natural one-to-one correspondence.*

If $f\colon W \to X$ and $g\colon Y \to Z$ are k-maps, then f and g induce functions $f^*\colon [X;Y] \to [W;Y]$, $g_*\colon [X;Y] \to [X;Z]$, respectively, where

$$f^*[h] = [hf], \quad g_*[h] = [gh], \quad h \in Y^X.$$

Indeed, it follows from (2.5) that

(2.7) *if $f \sim u\colon W \to X$ and $g \sim v\colon Y \to Z$, then $f^* = u^*$ and $g_* = v_*$.*

As before, S^q denotes the euclidean q-sphere with basepoint. If Y is any space, then since $S^q \times I$ is compact, every k-homotopy $F\colon S^q \times I \to Y$ is an ordinary homotopy; thus $\pi_q(Y) = [S^q;Y]$. In fact by (2.7), every k-homotopy equivalence $g\colon Y \sim Z$ induces isomorphisms $g_*\colon \pi_q(Y) \cong \pi_q(Z)$ in each dimension $q \geqq 0$.

Following Spanier [11, § 2] we say that a space Y with basepoint e is a *weak group* if there is a k-continuous multiplication function $m\colon Y \times Y \to Y$ such that the binary operation defined on Y by the formula

$$xy = m(x,y), \quad x,y \in Y,$$

gives Y the algebraic structure of a group with identity element e. Indeed, if G is an abelian group and if $K = |K(G,n)|$, the geometric realization of the Eilenberg-MacLane complex, then K is a weak abelian group [11, Theorem 3.6]; moreover, K is a CW-complex having a single nonvanishing homotopy group G in dimension n [7, Theorem 1 and Remark 2]. In fact, one can show that any two multiplication functions on K are k-homotopic when $n \geqq 1$. We also recall [8]:

(2.8) *$\Omega^q K \simeq |K(G, n-q)|$, $0 \leqq q \leqq n$, and $\Omega^q K$ is a contractible space if $q > n$.*

If Y is an arbitrary weak group if X is an arbitrary topological space, then it follows from (2.3) and (2.4) that the multiplication m on Y induces a k-continuous multiplication m^X on Y^X; hence m also induces a group structure on $[X;Y]$.

(2.9) *If Y is a weak abelian group, then Y^X is a weak abelian group and $[X;Y]$ is an abelian group.*

If Z is also a weak group and if the k-map $g: Y \to Z$ is a homomorphism, then the induced functions: $g^X: Y^X \to Z^X$ and $g_*: [X;Y] \to [X;Z]$ are likewise homomorphisms. From this fact and (2.7) we obtain:

(2.10) *If the k-map $g: Y \to Z$ is k-homotopic to a homomorphism and if g is a k-homotopy equivalence, then for all spaces X,*

$$g_*: [X;Y] \cong [X;Z].$$

Thus, for example, the homotopy equivalence $\Omega^q K \simeq |K(G, n-q)|$ of (2.8) is k-homotopic to a homomorphism since the multiplication on $|K(G, n-q)|$ is unique up to k-homotopy. Indeed, it follows from (2.13) below that for every pathwise connected space X, the multiplication on K^X is uniquely determined up to k-homotopy. We also note that if Y is a weak group, then the natural one-to-one correspondences of (2.6) are group isomorphisms with respect to the group structures induced by the multiplication on Y.

The *direct product* of the weak groups Y and Z is the weak group $Y \times Z$, i.e. the cartesian product space furnished with the usual multiplication. By (2.4) we have:

(2.11) *If Y and Z are weak groups, then the natural map $r: (Y \times Z)^X \to Y^X \times Z^X$ is an isomorphism and it induces an isomorphism $r_*[X; Y \times Z] \cong [X;Y] \times [X;Z]$.*

Let G be an abelian group and let $K = |K(G,n)|$, $n \geqq 0$. The group $[X^+;K]$ is called the *k-homotopical cohomology group of X in dimension n with coefficients in G.* If X is a CW-complex, then X is a k-space, so every k-map $h: X \to Y$ and every k-homotopy $H: X \times I \to Y$ is continuous. In this case, $[X^+;K]$ is the homotopical cohomology group and is isomorphic with the usual cohomology group of X [4]. An important generalization of this fact has been proved by P. J. Huber, using sheaf-theoretic techniques [5, fn. p. 76]:

(2.12) *If X is a paracompact Hausdorff space, then*

$$[X^+;K] \cong \check{H}^n(X;G).$$

Next, we recall a theorem of J. C. Moore [9, Theorem 3.29], [3, Satz 7.1]:

(2.13) *If Y is a pathwise connected weak abelian group, then*

$$Y \simeq \mathop{\times}_{q=1}^{\infty} |K(\pi_q(Y), q)|.$$

We conclude this section of preliminaries with a generalization of a theorem due to Thom [13]:

(2.14) *Let X be a paracompact space, let G be an abelian group and let $K = |K(G,n)|$, $n \geqq 0$. Then there is a k-homotopy equivalence*

$$K^{X^+} \sim \mathop{\times}_{q=0}^{n} |K(\check{H}^{n-q}(X;G),q)|.$$

Proof. Since K is a weak abelian group, then by (2.9), K^{X^+} is likewise a weak abelian group. Now let $L = |K(G,n+1)|$. Then L^{X^+} is also a weak abelian group; hence it follows from (2.13) that there is a homotopy equivalence:

$$(L^{X^+})_0 \simeq \mathop{\times}_{q=0}^{\infty} |K(\pi_{q+1}(L^{X^+}),q+1)|,$$

where $(L^{X^+})_0$ is the path component of the identity element of L^{X^+}. By (2.8) there is a homotopy equivalence $K \simeq \Omega L$; hence by (2.5) and (2.2),

$$K^{X^+} \sim (\Omega L)^{X^+} \simeq L^{X^+ \divideontimes S^1} \simeq \Omega(L^{X^+}) \times \Omega(L^{X^+})_0.$$

Thus,

$$K^{X^+} \sim \Omega \mathop{\times}_{q=0}^{\infty} |K(\pi_{q+1}(L^{X^+}),q+1)|.$$

But

$$\pi_{q+1}(L^{X^+}) \cong \check{H}^{n-q}(X;G), \qquad q = 0,1,\cdots,n.$$

For

$$\begin{aligned} \pi_{q+1}(L^{X^+}) = [S^{q+1};L^{X^+}] &\cong [X^+;L^{S^{q+1}}] \text{ by (2.6)}, \\ &\cong [X^+;|K(G,n-q)|] \text{ by (2.1) and (2.9)}, \\ &\cong \check{H}^{n-q}(X;G) \text{ by (2.12)}. \end{aligned}$$

Moreover, $\pi_{q+1}(L^{X^+}) = 0$ if $q+1 > n+1$. Consequently,

$$K^{X^+} \sim \Omega \mathop{\times}_{q=0}^{n} |K(\check{H}^{n-q}(X;G),q+1)| \simeq \mathop{\times}_{q=0}^{n} |K(\check{H}^{n-q}(X;G),q)|,$$

by (2.1), (2.4) and (2.9).

3. Proof of Theorem 1. Let $X \times Y$ be a paracompact space, let G be an abelian group and let $K = |K(G,n)|$. Then by (2.12),

$$\check{H}^n(X \times Y;G) \cong [(X \times Y)^+;K].$$

But clearly $(X \times Y)^+$ is homeomorphic with $X^+ \divideontimes Y^+$. Hence by (2.6),

$$\check{H}^n(X \times Y;G) \cong [X^+ \divideontimes Y^+;K] \cong [X^+;K^{Y^+}].$$

But X and Y are paracompact spaces since $X \times Y$ is paracompact; hence from (2.14),

$$K^{Y^+} \sim \bigtimes_{q=0}^{n} | K(\check{H}^{n-q}(Y;G), q)|.$$

So by (2.10), (2.11) and (2.12),

$$[X^+; K^{Y^+}] \cong \sum_{q=0}^{n} \check{H}^q(X; \check{H}^{n-q}(Y;G)).$$

Therefore,

$$\check{H}^n(X \times Y; G) \cong \sum_{q=0}^{n} \check{H}^q(X; \check{H}^{n-q}(Y;G)),$$

and Theorem 1 is proved.

4. Proof of Theorem 2. The universal coefficient theorem for cohomology proved by F. P. Peterson [10, Appendix] asserts that

$$0 \to \check{H}^q(X) \otimes G \to \check{H}^q(X;G) \to \check{H}^{q+1}(X) * G \to 0$$

is a natural exact sequence of groups, provided that either X is compact (G an arbitrary abelian group) or else X is paracompact and G is a finitely generated abelian group. Let L be a principal ideal domain and let G be an L-module. Then Peterson's argument also shows that

$$(4.1) \qquad 0 \to \check{H}^q(X;L) \otimes_L G \to \check{H}^q(X;G) \to \check{H}^{q+1}(X;L) *_L G \to 0$$

is a natural exact sequence of L-modules whenever X is compact and G is an arbitrary L-module or whenever X is paracompact and G is finitely generated over L.

Theorem 2 now follows from Theorem 1 on substituting $\check{H}^{n-q}(Y;L)$ for G on (4.1).

Added in proof: We find our homotopical method for obtaining a Künneth formula was anticipated by B. Eckmann and P. J. Huber in "Spectral sequences for homology and homotopy groups of maps," Battelle Research Report, Geneva, 1960, where the method was applied to a cartesian product of CW-complexes to yield our Theorem 1 for this case.

THE UNIVERSITY OF MICHIGAN.

REFERENCES.

[1] A. Borel, *Seminar on Transformation Groups*, Annals of Mathematics Studies No. 46, Princeton, 1960.

[2] R. Brown, "Function spaces and product topologies," *Quarterly Journal of Mathematics, Oxford* (2), vol. 15 (1964), pp. 238-250.

[3] A. Dold und R. Thom, "Quasifaserungen und unendliche symmetrische Produkte," *Annals of Mathematics* (2), vol. 67 (1958), pp. 239-281.

[4] B. Eckmann et P. Hilton, "Groupes d'homotopie et dualité: coefficients," *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris*, vol. 246 (1958), pp. 2991-2993.

[5] P. J. Huber, "Homotopical cohomology and Čech cohomology," *Mathematische Annalen*, vol. 144 (1961), pp. 73-76.

[6] J. L. Kelley, *General Topology*, New York, 1955.

[7] J. Milnor, "The geometric realization of a semi-simplical complex," *Annals of Mathematics* (2), vol. 65 (1957), pp. 357-362.

[8] ———, "On spaces having the homotopy type of a CW-complex," *Transactions of the American Mathematical Society*, vol. 90 (1959), pp. 272-280.

[9] J. C. Moore, "Semi-simplical complexes and Postnikov systems," *Symposium Internacional de Topología Algebraica*, Mexico City, 1958, pp. 232-247.

[10] F. P. Peterson, "Some results on cohomotopy groups," *American Journal of Mathematics*, vol. 78 (1956), pp. 243-258.

[11] E. Spanier, "Infinite symmetric products, function spaces, and duality," *Annals of Mathematics* (2), vol. 69 (1959), pp. 142-198.

[12] ———, "Quasi-topologies," *Duke Mathematical Journal*, vol. 30 (1963), pp. 1-14.

[13] R. Thom, "L'homologie des espaces fonctionnels," *Colloque de Topologie Algébrique*, Louvain, 1956, pp. 29-39.

[14] R. L. Wilder, *Topology of Manifolds*, American Mathematical Society Colloquium Publications, vol. 32, Providence, 1949.

FOLIATIONS AND PSEUDOGROUPS.

By RICHARD SACKSTEDER.*

1. **Introduction.** The main purpose of this paper is to describe some limitations on the topological properties of foliations of compact manifolds, especially foliations of co-dimension one. The discussion is in two parts. The first part is concerned with pseudogroups. The theorems given there were motivated by their applications to foliations, which are given in the second part; however, they may also be of some independent interest. The content of each part is described in more detail in a separate introduction.

Th author wishes to acknowledge numerous discussions with colleagues which were helpful in the preparation of this paper. He is especially indebted to Haefliger, Kadison, Lima, Rosenberg, A. J. Schwartz, and Smale.

Part I.

Pseudogroups.

2. **Introduction.** Van Kampen [6] has given an exposition of some of the classical theorems about groups of homeomorphisms of the 1-sphere S^1. Recently, A. J. Schwartz [17] has generalized one of these theorems, due to A. Denjoy [2], by considering a pseudogroup acting on a subset of the line $R = R^1$, rather than a group acting on S^1. The main motivation for these investigations has been to analyze the possible global behavior of trajectories of differential equations on 2-manifolds. G. Reeb [10] has suggested the investigation of analogous questions for foliations of co-dimension one of manifolds of arbitrary dimensions greater than two. This leads to certain questions about the possible actions of pseudogroups on S^1 and R which are more general than those discussed by Van Kampen and Schwartz, because one can no longer assume that the groups or pseudogroups are generated by a single element. Here we give more general theorems that are needed for applications to the study of foliations. These applications are discussed in Part II. Although most of the results here are concerned with pseudogroups

Received April 6, 1964.

* This work has been partially supported by the National Science Foundation under Grant NSF GP-1904.

which act on one-dimensional manifolds, Theorem 3, and its corollary are exceptions in which the pseudogroup acts on an arbitrary compact Hausdorff space and it is shown that a certain equicontinuity condition implies the existence of an invariant measure.

3. Pseudogroups acting on R^1. The first theorem is concerned with pseudogroups of diffeomorphisms which act on the real line, R. By a *pseudogroup* of homeomorphisms (diffeomorphisms, etc.) is meant a collection $\Gamma = \{f_a\colon a \in A\}$ of homeomorphisms (diffeomorphisms, etc.) of open subsets of a topological space S, which satisfies certain conditions. If D_a denotes the domain of f_a and R_a the range, the conditions can be stated in the following way:

(a) $f_a \in \Gamma$ implies that $f_a^{-1}\colon R_a \to D_a$ is in Γ.

(b) If $f_a, f_b \in \Gamma$, $D_a \cap D_b \neq \emptyset$, and $f\colon D_a \cup D_b \to R_a \cup R_b$ is a homeomorphism (diffeomorphism, etc.) such that $f(x) = f_a(x)$ when $x \in D_a$ and $f(x) = f_b(x)$ when $x \in D_b$, then $f \in \Gamma$.

(c) $f_a, f_b \in \Gamma$ implies that $f_a \circ f_b \in \Gamma$, where $f_a \circ f_b$ is defined for $x \in f_b^{-1}(R_b \cap D_a)$ by $f_a \circ f_b(x) = f_a(f_b(x))$.

(d) The identity map of S is in Γ.

A subset $S_0 \subset S$ will be called *invariant* under Γ if for every $x \in S_0$ and every $f_a \in \Gamma$ such that $x \in D_a$, $f_a(x) \in S_0$. A non-empty, compact, invariant subset $B \subset S$ is colled *minimal* if no nontrivial subset of B has all of these properties. The usual argument shows that every compact invariant set contains at least one minimal set. A compact, invariant, nowhere dense, perfect subset $S \subset R$ will be called *exceptional.* A subset $B \subset S$ is called P-stable if B is either finite or for every $x \in B$ there are points $f_a(x) \neq x$, $f_a \in \Gamma$, arbitrarily close to x. Minimal sets are P-stable, but not conversely.

A subcollection $\Gamma_0 \subset \Gamma$ is said to *generate* Γ if every element of Γ is generated from elements of Γ_0 by a succession of operations (a), (b), (c). If Γ_0 generates Γ and is finite, Γ is said to be *finitely generated.*

Theorem 1 can now be stated. Its proof is based on arguments used by Schwartz [17] although the fact that the pseudogroup is generated by more than one element produces some complications.

THEOREM 1. *Suppose that Γ is a pseudogroup of C^2 diffeomorphisms of R, and $C \subset R$ is a P-stable and exceptional subset. Suppose that Γ is finitely generated by Γ_0 and that the intersection of the domain of each element of Γ_0 with C is compact. Then there is an $x \in C$ and $f_a \in \Gamma$ such that*

$$f_a(x) = x \text{ and } |f'_a(x)| < 1. \tag{3.1}$$

According to the first statement of the conclusion (3.1), Theorem 1 is a kind of fixed point theorem. The second part of the statement asserts that not only is there a fixed point, but there is one which is very non-trivial. In applications to differential equations on 2-manifolds the existence of a fixed point suffices to prove the desired conclusion, the existence of a periodic orbit, but for applications to foliations of manifolds the non-triviality of the fixed point is required to give useful information.

It might be supposed that Theorem 1 is vacuous in the sense that the exceptional set C cannot exist. However, it is possible for exceptional sets to occur even for groups, cf. [14].

4. Proof of Theorem 1. The finiteness of Γ_0 and the condition on the domain of its elements imply that there is an open set D containing C with the property that there are positive constants M and θ such that for every $h \in \Gamma_0$ and x in the intersection of D with the domain of h,

$$M^{-1} \leqq |h'(x)| \leqq M, \qquad |h''(x)| \leqq \theta/M. \tag{4.1}$$

It can be supposed for the proof that Γ_0 and Γ are replaced by the collections which remain after the domain of each element is restricted to the intersection of its original domain with D. It will be supposed that Γ_0 is chosen in such a way that $h \in \Gamma_0$ implies $h^{-1} \in \Gamma_0$.

Now suppose that $h_1, \cdots, h_m$ are elements, not necessarily distinct, of Γ_0, and let $g_p = h_p \circ h_{p-1} \cdots \circ h_1$ for $1 \leqq p \leqq m$, where g_0 means the identity map on D. Then if the interval $[u, v]$ is in the domain of g_m,

$$|g'_m(u)| \leqq |g'_m(v)| \exp\left(\theta \textstyle\sum_{p=0}^{m-1} |g_p(u) - g_p(v)|\right). \tag{4.2}$$

This follows from

$$\begin{aligned} |\operatorname{Log} g'_m(u)/g'_m(v)| &\leqq \textstyle\sum_{p=1}^{m} |\operatorname{Log}|h'_p(g_{p-1}(u))| - \operatorname{Log}|h'_p(g_{p-1}(v))|| \\ &\leqq \textstyle\sum_{p=1}^{m} |h''_p(w_p)/h'_p(w_p)|\, |g_{p-1}(u) - g_{p-1}(v)| \\ &\leqq \theta \textstyle\sum_{p=1}^{m} |g_{p-1}(u) - g_{p-1}(v)|, \end{aligned}$$

where w_p is in the interval whose endpoints are $g_{p-1}(u)$ and $g_{p-1}(v)$. This proves (4.2).

Let q be an endpoint of a maximal open interval in R-C. Then if x is a point on the orbit $\Gamma_q = \{f_a(q) : f_a \in \Gamma\}$, we can assign an integer $k(x)$ as follows: We say that $k = k(x)$ if there are elements $h_1, h_2 \cdots h_k \in \Gamma_0$ (not necessarily distinct) such that $x = h_k \circ h_{k-1} \cdots \circ h_1(q)$, and $k = k(x)$ is the smallest integer for which this is possible. Let μ be a positive number so small that the interval of length μ centered at a point $y \in C$ lies completely in the domain of any element of Γ_0 whose domain contains y itself. Such a

number clearly exists. There are only finitely many intervals in R-C of length μ or more. Consequently, there is an integer $N \geqq 0$ such that every point $x \in \Gamma_q$ which is an endpoint of one of these intervals satisfies $k(x) < N$. The definition of N implies that if $x \in \Gamma_q$ and $k(x) \geqq N$, then the maximal open interval of R-C which has x as its endpoint is a subset of the domain of any element of Γ_0 whose domain contains x.

For any non-negative integer m, there are finitely many intervals of R-C which contain endpoints x such that $k(x) = m + N$. Let L_m be the length of the largest of these. Then if L is the length of an interval containing C, it is clear that

$$\sum_{j=0}^{m} L_j \leqq 2L. \tag{4.3}$$

Let $q_1, \cdots, q_t$ be the points of Γ_q satisfying $k(q_i) = N$ and let F, $F \leqq L_0 < \mu$ be the minimum of the lengths of the intervals of R-C containing one of the points q_i. Then if $x \in \Gamma_q$, $k(x) = m + N$, $m > 0$, there are elements $h_1, \cdots, h_m$ of Γ_0 such that $x = h_m \circ \cdots h_1(q_i) = g_m(q_i)$, for some $i = 1, \cdots, t$. It is clear that $x \neq h_j \circ \cdots h_1(q_i) = g_j(q_i)$ for any $j = 0, \cdots, m-1$, where by g_0 is meant the identity map on D. The main part of the proof consists in showing that there is a number ν, $0 < \nu < \mu$ such that for every $m = 0, 1, \cdots$ any map of the type g_m is defined on a ν-neighborhood of q_i, and for every point v in such a neighborhood

$$| g'_m(v) | \leqq K L_m, \tag{4.4}$$

where K and ν depend only on θ, L, and F.

For the moment, we suppose that g_m is defined on a subinterval $[q_i, v]$ of the interval of R-C containing q_i. Then if $0 \leqq j \leqq m$ and l_j is the length of $g_j([q_i, v])$ we have

$$L_j \geqq l_j \geqq | g'_j(w_j) | F, \tag{4.5}$$

where w_j is a point of $[q_i, v]$. Therefore, in view of (4.3),

$$\sum_{j=0}^{m} | g'_j(w_j) | \leqq 2L/F. \tag{4.6}$$

Now (4.2) and the definition of L imply

$$\begin{aligned} | g'_j(q_i) | &\leqq | g'_j(w_j) | \exp(\theta \textstyle\sum_{p=0}^{j-1} | g_p(q_i) - g_p(w_j) |) \\ &\leqq | g'_j(w_j) | \exp(2\theta L), \end{aligned} \tag{4.7}$$

hence by (4.6)

$$\sum_{j=0}^{m} | g'_j(q_i) | \leqq (2L/F) \exp(2\theta L) = \sigma. \tag{4.8}$$

The relation (4.8) holds independently of m and i.

Let $x_0 = q_i$ and let $x_j = g_j(x_0)$. Suppose $\lambda > 1$, let

$$\nu = \min(\mu(\lambda\sigma)^{-1}, (\theta\lambda\sigma)^{-1}\log\lambda),$$

and note that $\nu < \mu$ because of $\lambda\sigma > 1$. Now it will be proved by induction that if $|v - x_0| < \nu$, then $g_j(v)$ is defined,

$$(4.9)_j \qquad |g_j(v) - x_j| < \mu,$$

and

$$(4.10)_j \qquad |g'_j(v)| \leqq \lambda |g'_j(x_0)|.$$

The assertions $(4.9)_0$ and $(4.10)_0$ are obvious. The relation $(4.9)_{j-1}$ assures that $g_j(v) = (h_j \circ g_{j-1})(v)$ is defined for $|v - v_0| < \nu$. Then (4.2) gives

$$|g'_j(v)| \leqq |g'_j(x_0)| \exp(\theta \textstyle\sum_{p=0}^{j-1} |g_p(v) - g_p(x_0)|),$$

hence $(4.10)_p$ for $p = 0, \cdots, j-1$ and the mean value theorem give

$$|g'_j(v)| \leqq |g'_j(x_0)| \exp(\theta\lambda\nu \textstyle\sum_{p=0}^{j-1} |g'_p(x_0)|).$$

(4.8) now applies (where $q_i = x_0$), hence $\nu \leqq (\theta\lambda\sigma)^{-1}\log\lambda$ gives $(4.10)_j$. Therefore it has been shown that $(4.9)_{j-1}$ and $(4.10)_p$, $p = 0, \cdots, j-1$ imply $(4.10)_j$. If $(4.10)_j$ holds, the mean value theorem, (4.8), and $\nu \leqq \mu(\lambda\sigma)^{-1}$, give $(4.9)_j$. This completes the induction and shows that $(4.9)_j$ and $(4.10)_j$ hold for $j = 0, \cdots, m$. Note that ν is independent of m.

Now it follows that (4.4) holds for $|v - x_0| < \nu$, by $(4.10)_m$, (4.7), and (4.5), and K turns out to be $\sigma\lambda/2L$. To complete the proof of Theorem 1, note that there is a closed interval U with q in its interior such that there are elements $f_1, \cdots, f_t$ of Γ which are defined on U, satisfy $f_i(q) = q_i$, and $f_i(U) \subset \{v : |v - q_i| < \nu\}$. Since C is P-stable and exceptional, there are g_m for arbitrarily large m which map some q_i arbitrarily close to q. Also, (4.3) and (4.4) imply that $|(g_m \circ f_i)'(u)|$ can be made arbitrarily small uniformly for all $u \in U$. It follows that there is a g_m such that

$$(g_m \circ f_i)(U) \subset U \text{ and } |(g_m \circ f_i)'(u)| < c < 1$$

if $u \in U$.

It follows from these conditions that $f_a = g_m \circ f_i$ has a unique fixed point $x = \cap \{f_a{}^k(U) : k = 1, 2 \cdots\}$. Since U contains points of C and C is closed and invariant, $x \in C$. Therefore (3.1) holds and the proof is complete.

5. Germs which leave the origin fixed. Theorem 1 shows that in order for an exceptional set to exist, there must be an element of the pseudogroup Γ which has a non-trivial fixed point in the sense of (3.1). Theorem 2

shows that, on the other hand, it is in a certain sense not possible to have an exceptional set if there are too many non-trivial elements of Γ with the same fixed point.

THEOREM 2. *Let Γ be a pseudogroup of C^2 diffeomorphisms of R and suppose that for every $f_a \in \Gamma$, $f_a(0) = 0$. Suppose that the set $\{\mathrm{Log}\,|f'_a(0)| : f_a \in \Gamma\}$ contains two rationally independent numbers. Let $C \subset R$ be closed, invariant under Γ, and suppose that $0 \in C$. Then there is an $\epsilon > 0$ such that each of the open intervals $(-\epsilon, 0)$, $(0, \epsilon)$ is either a subset of C, or does not intersect C.*

If all of the maps in Γ are linear, Theorem 2 is essentially the Kronecker-Weyl Theorem, hence the theorem asserts that in a sense Γ acts in the same way as the linear parts of the elements of Γ act. Theorem 2 can be generalized to n-dimensions, but the statement is more complicated. The proof employs results of P. Hartman [5] instead of those of S. Sternberg [20], which we use here.

6. Proof of Theorem 2. It can be supposed that there is an element $g \in \Gamma$ such that $0 < g'(0) = a < 1$. The pseudogroup can be replaced by one whose elements are defined on the intersections of their original domain with some fixed open neighborhood. This implies that, according to a theorem of S. Sternberg [20], it can be supposed that $g(x) = ax$ on the domain where g is defined, provided that C^2 differentiability is reduced to C^1.

Now suppose that, for example, the interval $(0, \epsilon)$ contains points of C for arbitrarily small positive ϵ, but that for no ϵ is $(0, \epsilon) \subset C$. It will be shown that these conditions lead to a contradiction. If not, there must be a positive $x \in C$, with $(0, x) \subset$ domain g and an $f \in \Gamma$ such that f is defined on $[0, x]$, $f(t) = bt + o(|t|)$ where $0 < b < 1$, $0 < f(t) < t$ if $0 < t < x$, and $bx \notin C$. This is a consequence of the Kronecker-Weyl Theorem and the assumption that $\{\mathrm{Log}\,|f'_a(0)| : f_a \in \Gamma\}$ contains rationally independent numbers. Let $\delta > 0$ be so small that if $|bx - y| < \delta$, then $y \notin C$. For $k = 1, 2, \cdots$

$$|a^k bx - f(a^k x)| = o(|a^k x|) = a^k o(1).$$

Hence for large k, $|bx - a^{-k} f(a^k x)| < \delta$, which implies by definition of δ that

$$a^{-k} f(a^k x) \notin C. \tag{6.1}$$

But, since the domain of g contains $(0, x)$ and $0 < f(a^k x) < a^k x \leqq a^p x \leqq x$ for $p = 0, 1, \cdots, k$, $a^{-k} f(a^k x) = (g^{-k} \circ f \circ g^k)(x)$. But the invariance of C implies that this is a point of C, contradicting (6.1). This proves the desired statement about $(0, \epsilon)$. The statement about $(-\epsilon, 0)$ is proved in exactly the same way.

7. Functionals which are invariant under a pseudogroup. Let Γ be a pseudogroup which acts on a compact Hausdorff space S. Denote by $C(S)$ the space of real valued continuous functions on S with the uniform topology and let $P(S)$ denote the non-negative elements of $C(S)$. If $u \in C(S)$ and $g \in \Gamma$, the element $gu \in C(S)$ is defined whenever the range of g contains the support of u by $(gu)(x) = u(g(x))$ if x is in the domain of g and by $(gu)(x) = 0$ otherwise. A continuous linear functional $I\colon C(S) \to R$ is said to be invariant under Γ if $I(u) = I(gu)$ holds for all $g \in \Gamma$ and $u \in C(S)$ such that gu is defined. A linear functional I is said to be *positive* if $u \in P(S)$ implies $I(u) \geqq 0$. A positive linear functional I is said to be *strictly positive* if $I(u) = 0$ for $u \in P(S)$ only if $u = 0$. The pseudogroup Γ is said to be *equicontinuous at* $x \in S$, relative to some uniform structure $\{N_a(x)\colon a \in A, x \in S\}$ on S, if the following conditions hold: For every $a \in A$, there exist $b, c \in A$ such that if the domain of $g \in \Gamma$ contains $N_b(s)$ then $g(N_b(x)) \subset N_a(g(x))$ and if the domain of g contains $N_a(x)$, $g(N_a(x)) \supset N_c(g(x))$.

Theorem 3 is an analogue of a theorem of Segal [18], p. 114. Segal's theorem differs from Theorem 3 in two respects; it is valid for locally compact spaces S and for groups rather than for pseudogroups. It is probably not true that Theorem 3 (with (iii) replaced by $I_p \neq 0$) can be extended to locally compact spaces. Expressed in terms of measure rather than functionals, the difficulty seems to be that it is not natural for pseudogroups to require that the measure of every compact set be finite.

THEOREM 3. *Let Γ be a pseudogroup of local homeomorphisms of a compact space S. Then if Γ is equicontinuous at $p \in S$ there is a positive linear functional I_p with properties:* (i) *I_p is invariant under* Γ, (ii) *Support $I_p \subset C_p$, where C_p is the closure of the orbit Γ_p,* (iii) $I_p(1) = 1$.

Theorem 3 has the following corollary.

COROLLARY 7.1. *Assume that the hypotheses of Theorem 3 are satisfied for every $p \in S$. Suppose that there are countably many minimal sets $M_1, M_2, \cdots$ under Γ such that $\cup \{M_i\colon 1, 2, \cdots\}$ is dense in S. Then there is a continuous and strictly positive linear functional I which is invariant Γ and such that $I(1) = 1$.*

Remark. The proof is somewhat like Weil's construction of the Haar integral [21]. Some details are therefore omitted where the argument is essentially the same as Weil's. The main difficulty which arises in our proof and which is not present in proofs of the existence of Haar measure or in Segal's proof is that one cannot assume here that a compact set can be

covered by translates of arbitrarily small neighborhoods of a point, even if the orbit of the point is dense in the compact set. This situation arises because there is no assumption that the domains of elements of Γ are uniformly large. It is essential for the applications to foliations that this assumption not be made.

The following facts about uniform spaces will be needed.

LEMMA 7.1. *Let K be a compact subset of a uniform space S. Let $\{N_a(x) : a \in A, x \in S\}$ be a system of neighborhoods for S. Then for every $a \in A$, there is an integer $m = m(K,a)$ such that if $N_a(y_1), \cdots, N_a(y_r)$ is a set of disjoint neighborhoods with $N_a(y_i) \subset K$, then $r \leqq m$.*

Proof. Suppose that for every $i = 1, 2, \cdots$, there is a sequence $N_a(y_1{}^i), N_a(y_2{}^i), \cdots, N_a(y_i{}^i)$, of i disjoint neighborhoods, with $N_a(y_j{}^i) \subset K$. Let y_1 be a cluster point of the sequence $y_1{}^1, y_1{}^2, \cdots$. Let $b \in A$ be such that if $y \in N_b(x)$, then $z \in N_b(y)$, then $z \in N_a(x)$. One can then pick a subsequence of $y_1{}^1, y_1{}^2, \cdots, z_1{}^1, z_1{}^2, \cdots$ such that $z_1{}^i \in N_b(y_1)$, $i = 1, 2, \cdots$. Let $z_2{}^2, z_2{}^3, \cdots$ be the corresponding subsequence of $y_2{}^2, y_2{}^3, \cdots$. Then $z_2{}^2, z_2{}^3, \cdots$ has a subsequence which lies in $N_b(y_2)$ where y_2 is such that $N_b(y_1) \cap N_b(y_2) = \emptyset$. In fact, the choice of b shows that since $N_a(z_1{}^i) \cap N_a(z_2{}^i) = \emptyset$ and $z_1{}^i \in N_b(y_1)$ it is not possible that $N_a(z_2{}^i)$ contains a point of $N_b(y_1)$. One applies the definition of b taking $x = z_1{}^i$, $y = y_1$, $z =$ any point of $N_b(y_1)$.

By proceeding in this manner, one obtains a sequence $N_b(y_1), N_b(y_2), \cdots$ of disjoint neighborhoods. Clearly $y_i \in K$, hence the sequence $y_1, y_2, \cdots$ must have a cluster point $y_0 \in K$. However, this is easily seen to be impossible. For let c be such that if $y \in N_c(x)$ and $z \in N_c(y)$ then $z \in N_b(x)$. Taking $x = y_i$ where i is so large that $y_0 \in N_c(y_i)$, $z = y_j$ where $j \neq i$ and $y_j \in N_c(y_0)$, and $y = y_0$ contradicts $N_b(y_i) \cap N_c(y_j) = \emptyset$. This contradiction proves Lemma 7.1.

The proof of Theorem 3 depends on associating a non-negative number (u,v) to elements $u, v \in P(S)$ as follows: (u,v) is the least upper bound of all of the finite sums $\sum_{i=1}^n c_i$, where $c_i > 0$ and there exist $g_i \in \Gamma$ such that $g_i u$ is defined and $\sum_{i=1}^n c_i(g_i u)(x) \leqq v(x)$, for all $x \in S$. Here it is understood that $(u,v) = 0$ if there do not exist any c_i's and g_i's with the required properties. The lemma below shows that the definition makes sense and gives some properties of (u,v).

LEMMA 7.2. *Assume the hypotheses of Theorem 3. Then (u,v) as defined above has the following properties.*

(7.1) $\quad C(u,v) = (u, Cv) = (C^{-1}u, v)$,

(7.2) $\quad v \leqq w$ *implies* $(u,v) \leqq (u,w)$,

(7.3) $\quad (u,v)(v,w) \leqq (u,w)$,

(7.4) $\quad (u,v) + (u,w) \leqq (u, v+w)$,

(7.5) $\quad$ *If* gv *is defined,* $(u,v) = (u, gv)$,

(7.6) $\quad 0 \leqq (u,v) < \infty$ *if* $u(p) > 0$.

Only the proof of $(u,v) < \infty$ will be given because everything else is easy and is proved in practically the same way as in Weil's proof. In view of (7.1), it suffices to consider the case where $u(p) > 1$ and $\operatorname{Sup}\{v(x) : x \in S\} = 1$. Let u' be the characteristic function of the set where $u > 1$ and let v' be the characteristic function of the support of v. Then (u', v') can be defined just as (u,v) was defined for $u, v \in P(S)$, and it is clear that $(u',v') \geqq (u,v)$. Thus it suffices to prove that $(u', v') < \infty$. Let $N_a(p)$ be a neighborhood of $p \in S$ on which $u' = 1$. Applying the second equicontinuity property, there is a $c \in A$ such that $g(N_a(p)) \supset N_c(g(p))$. Let $d \in A$ be such that if $x_1 \in N_d(x_0)$, $x_2 \in N_d(x_1)$, then $x_2 \in N_c(x_0)$. Let $\{N_d(y_i) : i = 1, \cdots, m\}$ be a set of disjoint neighborhoods whose union is in the support of V and is such that the number, m, of neighborhoods is as large as possible, cf. Lemma 7.1.

Now it will be shown that $(u', v') \leqq m$. Let $c_1, \cdots, c_n, g_1, \cdots, g_n$ such that $c_i > 0$, $\sum_{i=1}^n c_i(g_i u')(x) \leqq v'(x) \leqq 1$. For each i, $g_i u' = 1$ on the set $N_c(g_i(p))$. We claim that for each i, there is a $j = j(i)$ such that $N_d(y_j) \subset N_c(g_i(p))$. This assertion follows from the definition of d by taking $x_0 = g_i(p)$, $x_1 \in N_d(x_0) \cap N_d(y_j)$ for some j, $x_2 =$ any point of $N_d(y_j)$. The existence of an index j such that $N_d(x_0) \cap N_d(y_j) \neq \emptyset$ is a consequence of the maximality of m. If u_i denotes the characteristic function of $N_d(y_j)$ when $j = j(i)$, it is clear that $\sum_{i=1}^n c_i u_i(x) \leqq 1$. Thus for any fixed j, $\sum_{i \in S(j)} c_i \leqq 1$, where $S(j)$ denotes the set of indices i such that $j = j(i)$. But then, $\sum_{i=1}^n c_i = \sum_{j=1}^m \sum_{i \in S(j)} c_i \leqq m$, hence $(u', v') \leqq m$ as asserted. This proves Lemma 7.2.

8. Proof of Theorem 3. First it will be proved that if $\epsilon > 0$ and $v, w \in P(S)$ are given, there is a neighborhood N of p such that if the support of u is contained by N and $u(p) > 0$, then

$$(u, v+w) \leqq (u,v) + (u,w) + \epsilon(u,1). \tag{8.1}$$

This will be verified first for the special case where $z = v + w$ never vanishes. Then $v' = v/z$ and $w' = w/z$ are continuous; hence the equicontinuity of Γ

at p implies that there is a neighborhood N_1 of p such that if $g \in \Gamma$ and the domain of g contains N_1, then both v' and w' vary by less than $a = \epsilon(z, 1)/6$ on $g(N_1)$. If $u \in P(S)$ and the support of u is contained in N_1, $u(p) > 0$, and $z(x) \geqq \sum_{i=1}^n c_i(g_i u)(x)$, then

$$v(x) = v'(x)z(x) \geqq \sum_{i=1}^n c_i(g_i u)(x)v'(x) \geqq \sum_{i=1}^n c_i(g_i u)(x)(v'(p_i) - a),$$

where $p_i = g_i^{-1}(p)$. Therefore, $(u, v) \geqq \sum_{i=1}^n c_i(v'(p_i) - a)$. Similarly, $(u, v) \geqq \sum_{i=1}^n c_i(w'(p_i) - a)$. It follows that

$$(u, v) + (u, w) \geqq (u, v + w)(1 - 2a). \tag{8.2}$$

The inequality (7.3) with the correspondences $u \to u$, $v \to v + w$, $w \to 1$, implies $(u, v + w) \leqq (u, 1)/(v + w, 1)$, hence (8.2) and $a = \epsilon(z, 1)/6$ give (8.1) for the case $z = v + w > 0$, where ϵ is replaced by $\epsilon/3$.

Now let the restriction that $z = v + w > 0$ be dropped. Applying the result already proved to the given v and with w replaced by $w(x) \equiv \epsilon$ gives if the support of u is in N_2 and $u(p) > 0$,

$$(u, v + \epsilon) \leqq (u, v) + 4\epsilon(u, 1)/3. \tag{8.3}$$

Similarly if the support of u is in N_3 and $u(p) > 0$

$$(u, w + \epsilon) \leqq (u, w) + 4\epsilon(u, 1)/3. \tag{8.4}$$

Also if the support of u is in N_4, and $u(p) > 0$,

$$(u, v + w + 2\epsilon) \leqq (u, v + \epsilon) + (u, w + \epsilon) + \epsilon(u, 1)/3. \tag{8.5}$$

Now (8.3), (8.4), and (8.5) give

$$(u, v + w + 2\epsilon) \leqq (u, v) + (u, w) + 3\epsilon(u, 1),$$

for u such that $u(p) > 0$ and whose support is in $N = N_2 \cap N_3 \cap N_4$. The desired result (8.1) then follows from (7.4) with the correspondence $v \to v + w$ and $w \to 2\epsilon$.

Now it can be shown that there exists a positive linear functional $I_p \colon C(S) \to R$ with the properties (i), (ii), (iii). This follows the Weil proof. Let for each $v \in P(S)$, S_v denote the interval $[0, (v, 1)^{-1}]$ and note that by (7.3) the value of $J_u(v) = (u, v)/(u, 1)$ lies in this interval. Let X be the Cartesian product of the spaces S_v and identify the function J_u with a point in this space in the obvious way. If N is any neighborhood of p, let $T(N)$ be the closure in X of the set of $J_u(v)$, where the support of u is in N and $u(p) > 0$.

Since X is compact and the sets $T(N)$ have the finite intersection

property, there is a point I_p in the intersection of all or the sets $T(N)$. One easily verifies that I_p is a positive continuous linear functional with the properties (i) and (iii). To see that I_p has the property (ii) note that by (iii) the support of I_p is non-empty and by (i) it is invariant under Γ. Thus if $v \in P(S)$ has its support in $K \subset S - C_p$ and N is a small enough neighborhood of p $J_u(v) = 0$ for every u whose support is a neighborhood of p contained in N. Therefore $I_p(v) = 0$ which proves (ii) and completes the proof Theorem 3.

To prove Corollary 7.1, one applies Theorem 3 to a sequence of points $p_1, p_2, \cdots$ where $p_i \in M_i$. One obtains a sequence of positive linear functionals $I_1, I_2, \cdots$. If $a_1, a_2, \cdots$ is a sequence of positive numbers such that $\sum_{i=1}^{\infty} a_i = 1$, it is easy to see that the functional $I = \sum_{i=1}^{\infty} a_i I_i$ has the desired properties. This proves Corollary 7.1.

9. Pseudogroups which act on S^1. A pseudogroup Γ of local homeomorphisms of a topological space S is said to *act freely* if, whenever $g \in \Gamma$ is such that $g(x) = x$ for some $x \in S$, then g is a restriction of the identity map.

LEMMA 9.1. *Let Γ be a pseudogroup of local homeomorphisms which acts freely on S^1. Suppose that every point $x \in S^1$ is contained in a minimal set. Then Γ is equicontinuous at every point $p \in S^1$.*

Proof. Let S^1 have an inner metric and let $N(\epsilon, p)$ denote the ϵ-neighborhood of $p \in S^1$. It must be shown that for every $p \in S^1$ and every $\epsilon > 0$ there are positive numbers δ_ϵ and λ_ϵ such that (i) if the domain of g contains $N(\delta_\epsilon, p)$, then $g(N(\delta_\epsilon, p)) \subset N(\epsilon, g(p))$, and (ii) if the domain of g contains $N(\epsilon, p)$, then $g(N(\epsilon, p)) \supset N(\lambda_\epsilon, g(p))$.

If (i) fails to hold for some p and ϵ, there is a sequence $g_1, g_2, \cdots$ of elements of Γ such that for large n, $g_n(N(1/n, p))$ contains a point q_n at exactly the distance ϵ from $g_n(p) = r_n$. It can be supposed $r = \lim r_n$ exists, and $\lim q_n = q$ exists. Since g_n^{-1} maps a set containing the arc (q_n, r_n) into $N(1/n, p)$, these conditions and the minimality assumption imply that the whole arc (q, r), which is of length ϵ, is in the closure of the orbit Γ_p. Let $g \in \Gamma$ and $a > 0$ be such that $g(N(a, p))$ is an arc whose endpoints are inside (q, r). Then for large n, $(q_n, r_n) \supset g(N(a, p))$, hence $g_n^{-1} \circ g$ is defined on all of $N(a, p)$ and maps $N(a, p)$ into $N(1/n, p)$. This implies that $g_n^{-1} \circ g$ has a fixed point x, hence $g_n^{-1} \circ g$ is a restriction of the identity, which is clearly a contradiction. This proves (i).

If (ii) fails to hold for some p and ϵ, there is a sequence $g_1, g_2, \cdots$ of elements of Γ such that if $r_n = g_n(p)$, $N(1/n, r_n)$ contains a point $q_n \notin g_n(N(\epsilon, p))$. It can be supposed that $q = \lim q_n = \lim r_n$ exists. This

implies that q is in the closure of the orbit of every point in some interval $[m,p] \subset N(\epsilon,p)$. The minimality condition then implies that there is a $g \in \Gamma$ and $a > 0$ such $g(N(a,q))$ is an interval with its endpoints inside $[m,p]$. Therefore, for some large n, $g^{-1} \circ g_n$ maps $[m,p]$ into itself, hence $g^{-1} \circ g_n$ has a fixed point. But this contradicts the assumption that Γ acts freely. This completes the proof.

THEOREM 4. *Let Γ be a pseudogroup of local homeomorphisms of S^1. Suppose that every point $p \in S^1$ is contained in a minimal set and that Γ acts freely on S^1. Then there is a strictly positive linear functional I which is invariant under Γ.*

Proof. This follows immediately from the fact that S^1 is the closure of the union of countably many minimal sets, Lemma 9.1, and Corollary 7.1.

Part II.

Foliations.

10. Introduction. Let V be an n-manifold with a foliated structure of co-dimension k, $(0 < k < n)$. The leaves of the foliation can be classified as follows: 1. compact leaves, 2. non-compact proper leaves, 3. leaves which are dense in V, 4. leaves which are dense in an open subset of V, but not in V itself, 5. exceptional leaves. Here, a leaf is called *proper* if its topology as a subset of V agrees with its topology as an $(n-k)$-manifold. An *exceptional* leaf F is one which is nowhere dense in V, but which has the property that every point $y \in F$ is in the closure in V of F-C for every compact subset $C \subset F$.

The main purpose of this part is relate the possibility of the existence of leaves of various types to holonomy properties of the leaves in the case where the foliation is of co-dimension one. Some general facts about holonomy of foliations of arbitrary co-dimension are discussed in Sections 11 and 12. The treatment here is rather sketchy beause most of these results are probably known. The remaining results are all concerned with foliations of co-dimension one. Theorem 5 in Section 13 shows that, under certain conditions, minimal exceptional leaves cannot exist. Theorem 6 in Section 14 shows that if the holonomy of the leaves is trivial, in a certain sense, the first real cohomology group of V is non-trivial. Finally, in Section 15, there are some theorems which describe special properties of foliations which come from a locally free action of R^{n-1} on an n-manifold.

11. Holonomy. Let V be an n-manifold with a C^2 foliated structure of co-dimension k, $0 < k < n$. Let F be a leaf of the foliation. The leaf F can be regarded as an $(n-k)$-manifold which is imbedded in $V, j: F \to V$. Suppose that one has a tubular neighborhood of the imbedding; that is, suppose that one has: (i) a vector bundle B over F, with fiber R^k, (ii) an open neighborhood Z of the zero cross-section $z: F \to B$, and (iii) an immersion $\phi: Z \to V$ such that $\phi \circ z = j$ and the images of the fibers of B intersect the leaves transversally. If $p, q \in F$, H_p and H_q will denote the fibers of B over p and q respectively. Corresponding to any curve g in F which begins at p and ends at q there is associated a C^2 diffeomorphism of a neighborhood of 0 in H_p onto a neighborhood of 0 in H_q (cf. [4], p. 380). The idea of the construction of this diffeomorphism is as follows: The foliation of V induces a foliation of B whose leaves intersect the fibers of B transversally. If F is identified with the zero cross-section of B, the curve g can be lifted along fibers to a curve g^* beginning at any point p^* of H_p sufficiently close to 0. The curve g^* will end at a point q^* in H_q, and the map $p^* \to q^*$ will be the desired diffeomorphism. Homotopic curves are associated to diffeomorphism which have the same germs. The set of germs of maps which arise in this way will be denoted by $G(p,q)$. If $p = q$, $G(p,p)$ is a group called the *Holonomy group at* p, which will also be denoted by $G(p)$. The correspondence between homotopy classes of curves and the elements of $G(p,q)$ depends on the tubular neighborhood, but only the following extent: If H'_p, H'_q, and $G'(p,q)$ are analogous to H_p, H_q, and $G(p,q)$ for some other tubular neighborhood, then there are germs of diffeomorphisms $f: H_p \to H'_p$ and $g: H'_q \to H_p$ such that if $h \in G(p,q)$ and $h \in G'(p,g)$ correspond to the same curve, then $h = g \circ h' \circ f$. If $p = q$, $g = f^{-1}$.

One can obtain a set of $k \times k$ matrices $L(p,q)$ (or $L(p)$) by differentiating the elements of $G(p,q)$ (or $G(p)$) at the origin. The correspondence between homotopy classes of curves and the elements of $L(p,q)$ (or $L(p)$) is well defined up to equivalences of the type which identify elements of $M \in L(p,q)$ (or $L(p)$) with elements TMS (or TMT^{-1}), where S and T are non-singular $k \times k$ matrices. The sets $L(p,q)$ define a parallel transport, hence a connection Ω on the normal bundle (or the associated frame bundle) of the immerision, $j: F \to V$. The connection thus defined has $L(p)$ for its holonomy group at p, and all of its local holonomy groups are trivial.

Some of the theorems stated below will have as an hypothesis that the group $L(p)$ is finite or trivial, so it is perhaps of interest to note that it is possible to determine whether this is the case by calculating the cohomology

class of a certain differential form if $k=1$. This is seen as follows: Let Ω be the connection described above on the frame bundle B associated to the normal bundle of the immersion $j: F \rightarrow V$ and let $v: B \rightarrow R$ define a volume on the fibers of B, cf. appendix. Part (ii) of Proposition A. 1 of the appendix shows the form $\theta \mid F$ associated to v and Ω is a cocycle in F. Moreover, part (i) of the same proposition shows that there is a volume invariant under Ω if and only if $\theta \mid F$ is a coboundary. A Riemannian metric on V induces a map $v: B \rightarrow R$ which defines a volume on the fibers of B. There is a short discussion in [15] showing how the form $\theta \mid F$ can be computed from the forms $\omega^1, \cdots, \omega^k$ which define the foliation locally and their exterior derivatives in this case. Thus the cohomology class of $\theta \mid F$ is, in principle, always computable by rather direct methods. In the case where $k=1$, it is clear that it is possible to have a volume invariant under Ω if and only if the linear holonomy group $L(p)$ is finite. This justifies the assertion made above.

12. Holonomy of leaves whose closure is minimal. If V is a foliated manifold, a subset $C \subset V$ is said to be *saturated* (or *invariant*) if it contains the leaf through each of its points. A compact, saturated subset $C \subset V$ is called *minimal* if C has no non-trivial subsets with these properties. The usual proof shows that every compact saturated subset contains a minimal set cf. [9], Proposition 2.

LEMMA 12.1. *Let V have a foliated structure of co-dimension one and let C be a minimal subset of V which is not all of V. Suppose that A is a smooth curve which intersects the leaves of the foliation transversally and whose endpoints are not in C. Then there is a pseudogroup Γ of local diffeomorphisms of a neighborhood of A which is finitely generated and which contains repersentatives for all of the germs of diffeomorphisms in $G(p,q)$ for all $p, q \in B = A \cap C$. In particular, if $p \in B$, $G(p)$ is a finitely generated group.*

Remark. One might suspect that the lemma is true because the fundamental group of each leaf in C is finitely generated. However, leaves satisfying the conditions of those in C in Lemma 12.1 can have fundamental groups which are not finitely generated as an example in [14] shows.

Proof of Lemma 12.1. Suppose that V has a Riemannian metric. Let $p \in C$. Then there exists an open neighborhood U of p with the following properties: (i) Each component (or plaque) of the intersection of a leaf with U is geodesically convex with respect to the metric induced on the leaves. (ii) There is a smooth transversal curve A_p passing through p which

intersects each plaque transversally and exactly once, and whose endpoints are not in C. The proof of this fact is tedious, but not difficult. The argument is similar to that used to prove the existence of geodesically convex neighborhoods of a point of a Riemannian manifold. We omit the proof here. A finite number of neighborhoods of the type described cover C. Denote them by $U(1),\cdots,U(k)$, let $A(1),\cdots,A(k)$ be corresponding transversal curves, and $B(i)=A(i)\cap C$. It is sufficient to prove the theorem in the case where $A(1)=A$, as will be shown at the end of the proof. Thus assume that $A(1)=A$ for the present.

Let $g=\{g(t):0\leqq t\leqq 1\}$ be a smooth curve on a leaf F, and suppose that $g(t)\subset S=\cup\{U(i):1,\cdots,k\}$. It is not assumed that $F\subset S$, necessarily. An ordered set of integers $i_1,\cdots,i_r$ will be called a *visitation sequence* for g if $1\leqq i_j\leqq k$ and there are $t_0=0<t_1<\cdots<t_r=1$ such that

(12.1) $$t_{j-1}\leqq t<t_j \text{ implies } g(t)\in U(i_j) \text{ for } j=1,\cdots,r$$

and

(12.2) $$g(t_j)\notin U(i_j) \text{ for } j=1,\cdots,r-1.$$

A curve can clearly have more than one visitation sequence.

Now let g_1 and g_2 be curves in the same leaf F with the same visitation sequence, $i_1,\cdots,i_r$, the same initial point $g_1(0)=g_2(0)$, and the properties that $g_1(1)\in A(i_r)$ and $g_2(1)\in A(i_r)$. It will be proved that

(12.3) $$g_1(1)=g_2(1) \text{ and } g_1 \text{ and } g_2 \text{ are homotopic in } F.$$

First observe that the property (i) implies (12.3) for the case $r=1$. The same property also shows that if $h_1=\{h_1(t):0\leqq t\leqq s_1\}$ and $h_2=\{h(t):0\leqq t\leqq s_2\}$ are two curves in F with the same visitation sequence i, j of length 2, and the same initial point $h_1(0)=h_2(0)$, then h_2 can be deformed within $F\cap\{U(i)\cup U(j)\}$ to the curve formed by h_1 followed by the geodesic in F connecting $h_1(s_1)$ to $h_2(s_2)$. To prove (12.3) in general, let s_1 (or s_2) be the parameter value for g_1 (or g_2) corresponding to t_1 in (12.1). Let $h_1=\{g_1(t):0\leqq t\leqq s_1\}$ and $h_2=\{g_2(t):0\leqq t\leqq s_2\}$. Applying the remark just made, one sees that g_2 is homotopic in F to a curve g_3 which consists of three pieces joined together as follows: The first piece is the curve h_1, the second piece is the geodesic in F connecting $g_1(s_1)$ to $g_2(s_2)$, and the third coincides with the subarc of g_2 defined by $s_2\leqq t\leqq 1$. It is clear that (12.3) holds if and only if it holds with g_2 replaced by g_3. Since g_1 and g_3 agree up to a point corresponding to a parameter value in

the second term of their visitation sequence, the number of terms in the visitation sequence is effectively one less. Thus (12.3) is proved by induction on r.

Let $i_1, \cdots, i_r$ be a finite ordered set of integers such that $1 \leqq i_j \leqq k$. Then (12.3) shows that to any such sequence one can associate a unique local diffeomorphism of an open (but possibly empty) subset of $A(i_1)$ to an open subset of $A(i_r)$ by associating to each point $A(i_1)$ the point $p \in A(i_r)$ (if any) which can be reached by curve g which is in the set S and in the leaf containing p, and which has the visitation sequence $i_1, \cdots, i_r$. The restriction of such a local diffeomorphism to a maximal subinterval in its domain is a diffeomorphism, and a finite number of these intervals cover $B(i_1)$. Thus to each sequence there correspond a finite number of diffeomorphisms. A sequence will be called a *simple* if $i_j \neq i_m$ unless $j = m$, and will be called a *simple cycle* at i_1 if $i_1 = i_r$ but $i_j \neq i_m$ in every other case where $j \neq m$. Clearly there exist only a finite number of simple sequences and simple cycles.

Now let $x \in B(i)$ for some $i = 2, \cdots, k$. Since C is minimal, one can find a curve g in the leaf containing x which begins at x and ends at a point of $B(1)$. Any such curve induces a diffeomorphism of an open interval of $A(i)$ containing x to a subinterval of $A(1)$. It can be supposed, since C is nowhere dense, that the endpoints of these intervals are not in C. A finite number of open intervals of this type cover $B(i)$ and it can be supposed that this cover consists of disjoint open intervals. The set of diffeomorphisms which arise in this way for $i = 2, \cdots, k$ is finite, and each point of $\cup \{B(i):$ $i = 2, \cdots, k\}$ is in the domain of exactly one of them. We denote these diffeomorphisms by $f_1, \cdots, f_q$, where we include the identity map of a neighborhood of $A(1)$ among the f_i.

Now it will be shown that the pseudogroup of local diffeomorphisms of $A(1)$ which is generated by diffeomorphisms of the type $f_a \circ h \circ f_b^{-1}$, where h corresponds to a simple sequence or a simple cycle, contains representatives of all of the germs in $G(p, q)$ for $p, q \in B(1)$. The diffeomorphisms $f_a \circ h \circ f_b^{-1}$ should be understood to be defined on the largest set where the composition makes sense and of course either f_a or f_b can be the identity map. Suppose that a visitation sequence for a curve which induces a germ in $L(p, q)$ is $i_1 = 1, i_2, \cdots, i_r$. If the curve has no repetitions other than possibly $i_1 = i_r$ $= 1$, it is a simple sequence or cycle and there is nothing to show. Otherwise, suppose that the first repetition encountered is $i_j = i_m$, where $1 \leqq j < m \leqq r$. Then $i_j, \cdots, i_m$ is a simple cycle. Let g be modified as follows: change the part of g corresponding to the term i_j in its visitation sequence so that g

intersects $A(i_j)$ in y. Join y to a point $z = f_c(y) \in B(1)$ by a curve which induces f_c. Then return to y along the same curve. Then continue along g as before. The part of the curve from p to z corresponds to a diffeomorphism of the type $f_c \circ h_1$, where h_1 corresponds to a simple sequence. The part from z to a point of $U(i_m)$, which can be assumed to be on $B(i_m)$, will induce a diffeomorphism of the type $h_2 \circ f_c^{-1}$ where h_2 corresponds to a simple cycle. Now the same process can be continued. It will end after a finite number of steps and thus the diffeomorphism corresponding to g has the same germ as one of the type generated by elements of the form $f_a \circ h \circ f_b^{-1}$, as was asserted. This completes the proof, for the case $A(1) = A$, because there are only finitely many elements of this type.

To prove the lemma in general all that needs to be done is to construct a finite number of diffeomorphisms from open intervals of A to $A(1)$ such that $B = A \cap C$ is covered by the union of the domains. Then each germ in $G(p, q)$ for $p, q \in B$ will correspond to a germ $G(p', q')$ for some $p', q' \in B(1)$, and the generators for $G(p', q')$ will induce generators for $G(p, q)$. The homomorphisms can be constructed as those from $A(i)$ to $A(1)$ were constructed. This proves the lemma.

13. Foliations with trivial linear holonomy. For the rest of this paper V will be an n-manifold with a C^2-foliated structure of co-dimension one. The linear holonomy group $L(p)$ for $p \in V$ will be said to have *rational dimension* m if m is the dimension of the vector space generated over the rational numbers by $\{\log |a| : a \in L(p)\}$. In view of the discusssion in Section 11, it is clear that an equivalent condition is that the 1-form $\theta \mid F$ has exactly m rationally independent periods. In particular, $L(p)$ has rational dimension 0 if $H^1(F, R) = 0$, or more generally if $\theta \mid F$ cohomologous to 0, where F is the leaf containing p.

The following theorem is perhaps our main result. It is a generalization of a theorem of A. J. Schwartz [17].

THEOREM 5. *Let $C \subset V$ be a nowhere dense minimal set. Suppose that for every $p \in C$, the rational dimension of $L(p)$ is different from one. Then C is a compact leaf.*

The proof of Theorem 5 is just a matter of putting together results which have already been proved. Since C is nowhere dense, there is a smooth curve A passing through any point $p \in C$ which intersects the leaves transversally and is such that the endpoints of A are not in C. Lemma 12.1 shows that there is a pseudogroup Γ of diffeomorphisms of open subsets of A which is finitely generated and contains representatives for all of the germs

in $G(p_1, p_2)$ for $p_1, p_2 \in C \cap A$. The minimality of C implies the minimality of $C \cap A$ under Γ. If C is not a compact leaf all of the leaves in C are exceptional. The latter condition implies that $C \cap A$ is exceptional in the sense of Theorem 1. Thus Theorem 1 shows that for some $p \in A \cap C$, the rational dimension of $L(p)$ is at least one. But Theorem 2 shows that the dimension is at most one. This proves Theorem 5.

14. Bundle-like metrics. A manifold V with a foliation of co-dimension one will be said to have a *bundle-like metric* if there is a Riemannian metric on V such that the distance between nearby leaves measured along orthogonal trajectories to the leaves is constant "locally." More precisely, what is meant is the following: For every $p \in V$, there is a neighborhood U and a homeomorphism $f\colon U \to S = \{(x^1, \cdots, x^n) : 0 < x^i < 1\}$ such that the subset of S where $x^n = c$ is a plaque and the lines $x^i = a^i$ $i = 1, \cdots, n-1$ are the isometric images of pieces of orthogonal trajectories, taking the obvious metric in S, cf. Reinhart [12].

LEMMA 14.1.[1] *If V is compact and admits a bundle-like metric, then*

(i) *all of the holonomy groups $G(p)$ are finite and*

(ii) $H^1(V, R) \neq 0$.

Proof. The conclusion (i) is obvious. In fact, the representatives of the germs in $G(p)$ must be local isometries, hence $G(p)$ contains, at most, two elements.

If the foliation is not oriented, let V be replaced by a two sheeted covering V', such that the foliation induced on V' is oriented. Since $H^1(V, R) \neq 0$ if $H^1(V', R) \neq 0$, it suffices to prove the latter. Let the 1-form ω define foliation and be of length one with respect to metric on 1-forms induced by the Riemannian metric on V'. These conditions define a form ω globally. With the local coordinates defined by neighborhoods of the type U, $\omega = dx^n$. Clearly, $d\omega = 0$, hence ω represents a cohomology class in $H^1(V, R)$. Since V' is compact and ω never vanishes, it is not possible that ω is cohomologous to 0. Therefore $H^1(V', R) \neq 0$ and $H^1(V, R) \neq 0$.

The following theorem is a converse of the conclusion of (i) of Lemma 14.1. It is related to a theorem of Haefliger [4], p. 390 or [3], p. 317.

THEOREM 6. *Let V be a compact connected manifold with a C^2-foliation of co-dimension one and suppose that all of the holonomy groups $G(p)$ for $p \in V$ are finite. Then there is a bundle-like metric on V and $H^1(V, R) \neq 0$.*

[1] (*added in proof*): A somewhat stronger result is proved by Reeb [11], p. 110. The assertion $H^1(V, R) \neq 0$ in the conclusion of Theorems 6 and 9 can be strengthened accordingly.

Proof. Standard arguments cf. [19] show that it is possible to construct a closed curve, S^1, whose intersections with leaves are all transversal. Moreover, it can be supposed that there is a Riemannian metric on V such that the intersections of S^1 with leaves are orthogonal. This can be accomplished by simple modifications in one of the standard proofs of the existence of Riemannian metric. The foliation induces a pseudogroup Γ of orientation-preserving local diffeomorphisms of S^1. The hypothesis on the holonomy group implies that Γ acts freely on S^1.

Now it will be shown that there is a metric on S^1 which is invariant under Γ. Theorem 1 implies that every minimal set is either dense or finite, hence if every point of S^1 is in a minimal set, the desired metric exists by Theorem 4. Suppose, if possible, that there is a point $p \in S^1$ which is not in a minimal set. Then there is a point q in the closure of its orbit which lies on a finite orbit. In this case, the leaf through p contains the leaf through q in its closure. The leaf through q must be proper, because the orbit through q is finite. But these conditions contradict Theorem 1 of [16]. This shows that S^1 has a metric which is invariant under Γ.

Now the metric on V will be modified to have the desired properties. The new metric will leave the length of tangent vectors to the leaves unchanged; however, the distances measured along the orthogonal trajectories will be modified. Along the orthogonal trajectory S^1 distance will be measured by the metric which is invariant under Γ. If $p \in S^1$ is connected to a point q by a curve g which lies in a leaf, the distance along the orthogonal trajectory through q for points near q is measured by the distance induced from the metric on S^1 by the representative of $G(p,q)$ which corresponds to g. The invariance of the metric on S^1 under Γ implies that the metric defined locally on a part of the orthogonal trajectory through q is independent of the choice of the curve g. Every point $q \in V$ lies on a leaf which intersects S^1 as is proved in [16], Theorem 4. This fact depends, of course, on the assumption that the holonomy groups are finite. Therefore a metric has been defined in all of the orthogonal trajectories in V.

The metric which has been defined on the orthogonal trajectories in V need not be differentiable with respect to the original atlas on V. Nevertheless, one can define neighborhoods U as described in the beginning of this section. First map a neighborhood $U \subset V$ to $S = \{(x^1, \cdots, x^n) : 0 < x^i < 1\}$ by a diffeomorphism f such that each set $x^n = c$ is the image of points in a single leaf and each line $x^i = a^i$, $i = 1, \cdots, n-1$ is the image of a subset of an orthogonal trajectory. The define a homeomorphism $h: S \to S$ such that $h(x^1, \cdots, x^{n-1}, x^n) = (x^1, \cdots, x^{n-1}, a\emptyset(x^n) + b)$ where $\emptyset(x^n)$ is the distance of $(x^1, \cdots, x^n)$ from the point $(x^1, \cdots, x^{n-1}, 1/2)$ along the orthogonal tra-

jectory and a and b are constants chosen in such a way that h becomes a homeomorphism onto S. Letting $f = h \circ f$ produces a map of the required type. This proves that V has a bundle-like metric. For now the metric on orthogonal trajectories is differentiable with respect to the local coordinates $(x^1, \cdots, x^n)$ defined by $f: U \to S$, thus a length is assigned to tangent vectors (relative to these coordinates) to the orthogonal trajectories. Thus, together with the length already assigned to tangent vectors to leaves, defines a Riemann metric on V with respect to the atlas defined on V by the local coordinates $f: U \to S$. This proves that V has a bundle-like metric. The rest of the conclusion follows from Lemma 14.1.

15. Locally free actions of R^{n-1}.[3] A foliation of co-dimension $n-k$ of an n-manifold V is said to be defined by a *locally free action* of R^k if there is a C^2-action $R^k \times V \to V$ whose orbits are the leaves of the foliation. Such an action defines a family $X_1, \cdots, X_k$ of vector fields which commute (i.e. $[X_i, X_j] = 0$) and are linearly independent at each point (cf. [7]). Foliations of this type have been investigated by Lima and Rosenberg [7], [13].

THEOREM 7. *Suppose that a foliation of the n-manifold V is defined by a locally free action of R^{n-1}. Suppose that F is a leaf whose linear holonomy group contains more than two elements. Then F is a proper leaf.*

THEOREM 8. *Suppose that a foliation of the n-manifold V is defined by a locally free action of R^{n-1}. Then there are no exceptional leaves whose closure is minimal.*

THEOREM 9. *Suppose that V is a compact n-manifold and a foliation of V is defined by a locally free action of R^{n-1}. Suppose that no leaf of V is compact. Then all of the leaves are dense in V, there is a bundle-like metric on V, and $H^1(V, R) \neq 0$.*

Proof of Theorem 7. Parts of this proof were suggested by arguments of Lima [7]. The assumption on the holonomy group of F implies that there is a cohomology class $C \in H^1(F, R)$ such that $0 < C(g) < 1$, where g is a closed curve in F. It can be supposed that g is an integral curve of the vector field $X = \sum_{i=1}^{n-1} a_i X_i$, where the a_i are constants. Let $p \in g$. If F is not proper, there is sequence $p_1, p_2, \cdots$ on the orthogonal trajectory to the leaves through p such that $p_i \in F$ and $p = \lim p_i$. A point $x = (x^1, \cdots, x^{n-1}) \in R^{n-1}$ acts on F by sending $q \in F$ to the point xq at "time" one away from q along the trajectory of the vector field $\sum_{i=1}^{n-1} x^i X_i$. In particular, $p_i = T_i p$

[3] (*added in proof*): The theorems and proofs in this section actually valid for actions $\emptyset: P^k \times V \to V$ whose orbits are all of dimension $n-1$, even if $k \neq n-1$. ($k = n-1$ is required if the action is to be locally free.)

for some $T_i \in R^{n-1}$. The vector field X is invariant under the action of R^{n-1}, hence T_i maps the curve g to the closed integral curve g_i of X passing through p_i. Moreover, the continuity of X implies that if $g(t)$ (or $g_i(t)$) is the point of g (or g_i) at time t away from p (or p_i), then

$$(15.1) \qquad \lim g_i(t) = g(t) \text{ uniformly in } t.$$

It is easy to see that there is an imbedding of $M = S^1 \times R^{n-2} \times (-\epsilon, \epsilon)$ onto an open subset of V with the properties: (i) p is the image of $(0,0,0) \in M$, (ii) $S^1 \times (0,0)$ is the preimage of g. (iii) For each fixed $(s,y) \in S^1 \times (-\epsilon, \epsilon)$, the image of the set $\{(s,x,y) : (s,x,y) \in M\}$ is a subset of a single leaf. (iv) For fixed $(s,x) \in S^1 \times R^{n-2}$, the image of $\{(s,x,y) : (s,x,y) \in M\}$ is a subset of an orthogonal trajectory. The image of M contains the curves g_i for large i by (i), (ii), and (15.1). For such i the preimage can be represented by a curve $\{(s_i(t), x_i(t), y_i(t)) : 0 \leqq t \leqq 1\}$ that $g_i(t)$ is the image of $(s_i(t), x_i(t), y_i(t))$. The curve $h_i = \{s_i(t), 0, y_i(t)) : 0 \leqq t \leqq 1\}$ is a closed curve, which is on the preimage of a single leaf, by (iii). The conditions (iii) and (iv) imply that the preimage of each leaf intersects the set $\{(s,0,y) : (s,0,y) \in M\}$ transversally, hence each component of the intersection must be a curve. Since the curves h_i are closed and $\lim y_i(0) = 0$, there must be components which are closed arbitrarily close to the preimage of g. However, this is impossible because the condition $0 < C(g) < 1$ implies that every component which contains a point close enough to $(0,0,0)$ must spiral in to the preimage of g. This proves Theorem 7.

Proof of Theorem 8. Suppose that there is an exceptional leaf F with minimal closure. Then Theorem 5 implies that some leaf in the closure of F has a linear holonomy group with infinitely many elements. But then Theorem 7 shows that the latter leaf is proper, hence compact by the minimality assumption. But then the leaf F must be compact, hence not exceptional. This proves Theorem 8.

Proof of Theorem 9. The conclusion that all leaves are dense follows from Theorem 8 and the known result that the conclusion of Theorem 8 implies that every non-dense leaf contains a compact leaf in its closure, cf. [9], [11]. Then if there is a leaf with a holonomy group which contains more than two elements, it is easy to see that there is a possibly different leaf F whose holonomy group contains a germ represented by a diffeomorphism which satisfies $0 < f(x) < x$ for sufficiently small positive x. Since F is dense, it contains points which correspond to such x. But then the proof of Theorem 7 shows that F is proper. This is impossible, hence every holonomy group must have, at most, two elements. But then Theorem 6 applies and gives the remaining conclusions.

Appendix.

Invariant Volumes on the Fibers of a Bundle.

The material in this section is probably known, but there does not seem to be a reference which is suitable for our purposes.

Let $G = GL(k, R)$ be the group of non-singuar $k \times k$ matrices. Let B be a C^1 principal bundle over a manifold F, with G as its group and fiber, and the projection $\pi\colon B \to F$. Let $\phi\colon G \to R$ be defined by $\phi(g) = \operatorname{Log} |\det(g)|$ for $g \in G$. The map ϕ is an analytic homomorphism. A C^1 map $v\colon B \to R$ will be said to define a *volume on the fibers* of B if for any local trivialization $\psi\colon G \times U \to \pi^{-1}(U)$, $(v \circ \psi)(g, u) = c(u) + \phi(g)$, where $(g, u) \in G \times U$. The function $c\colon U \to R$ depends, in general, on the trivialization.

Now suppose that there is a connection Ω (cf. [1] or [8]) on B as well as a volume on the fibers of B defined by a map v. It will be shown that there is associated to Ω and v a 1-form θ on F. Let ψ be a trivialization as above, and let X be a tangent vector to F at $u \in U$. Denote by $h(X, g)$ the unique horizontal tangent vector to B at $b = \psi(g, u)$ such that $(\pi_*)(h(X, g)) = X$. Then $(\psi^{-1})_*(h(X, g)) = X + Y(X, g)$, where $Y(X, g)$ is a tangent vector to G at g. The map $X \to Y(X, g)$ is linear. (In classical notation if $X = (dx^1/dt, \cdots)$, $g = (g_j{}^i)$, then $Y = dg_j{}^i/dt = \Gamma_{rs}{}^i g_j{}^r dx^s/dt$.) If e is the identity in G, the invariance of the horizontal subspace under the action of G implies that

$$Y(X, g) = Y(X, e)g. \tag{A.1}$$

LEMMA A.1. (Liouville's formula) *For every* $g \in G$

$$\langle d\phi(g), Y(X, g)\rangle = \operatorname{Trace} Y(X, e). \tag{A.2}$$

Proof. A direct computation shows that if $g = (g_j{}^i)$, then $(\partial\phi/\partial g_j{}^i)(g) = G_j{}^i$, where $(G_j{}^i)$ is the transpose of g^{-1}. If $m = (m_j{}^i)$ is any $k \times k$ matrix, $\sum_{ij} G_j{}^i m_j{}^i = \operatorname{Trace}(g^{-1}m)$. Taking $m = Y(X, g)$ and using (A.1) gives (A.2).

Now let γ be a smooth curve in U beginning at p and ending at q. Let γ^* be the horizontal life of γ to B which begins at $p^* \in \pi^{-1}(p)$ and ends at $q^* \in \pi^{-1}(q)$.

LEMMA A.2. *The map* $X \to \langle dc(p), X\rangle + \operatorname{Trace} Y(X, e)$ *defines a 1-form* θ_U *on* U. *Moreover,*

$$\int_\gamma \theta_U = v(q^*) - v(p^*), \tag{A.3}$$

hence θ_U *is independent of the trivialization.*

Proof. The first statement is obvious from the linearity of $Y(X, e)$ and the trace. To check (A. 3), note that if $p^*(t)$, $0 \leqq t \leqq 1$ is a parametrization of γ^*, and $p(t) = \pi \circ p^*(t)$ is a corresponding parametrization of γ, then there is a curve $g(t)$, $0 \leqq t \leqq 1$ in G such that $\psi(g(t), p(t)) = p^*(t)$. It follows that $v(p^*(t)) = c(p(t)) + \phi(g(t))$. Differentiating with respect to t gives $dv(p^*(t))/dt = \langle dc(p(t)), X_t \rangle + \langle d\phi(g(t)), g'(t) \rangle$, where X_t is the tangent to $p(t)$. The definition of a horizontal lift implies that the last term is equal to $\langle d\phi(g(t)), Y(X_t, g(t)) \rangle$, hence (A. 2) gives $dv(p^*(t))/dt = \langle dc(p(t)), X_t \rangle + \operatorname{Trace} Y(X_t, e)$, which is essentially (A. 3). In classical notation, $\langle \theta_U, X \rangle = \sum_{i,j} \Gamma_{ij}{}^i(x) dx^j/dt$. The next lemma is an immediate consequence of Lemma A. 2.

LEMMA A. 3. *There is a unique* 1-*form* θ *in* F *which agrees with* θ_U *for every trivialization* $\psi\colon G \times U \to \pi^{-1}(U)$ *of an open set* $U \subset F$. *If* γ *is a smooth curve in* L *and* γ^* *is the horizontal lift of* γ *which connects* p^* *to* q^*, *then*

$$\int_\gamma \theta = v(q^*) - v(p^*). \tag{A. 4}$$

The 1-form θ of Lemma A. 3 depends on both the connection and the volume defined by v. If v_1 and v_2 are two functions which define volumes on the fibers of B, and θ_1 and θ_2 are the corresponding 1-forms, it is clear that $\theta_1 = \theta_2 + df$, where f is defined globally on F. Thus the properties that θ is closed or exact depend only on Ω and not on v. The volume defined by v will be said to be invariant under Ω if for every curve γ and horizontal lift γ^* of γ (as in Lemma A. 3), $v(p^*) = v(p)$. It follows from (A. 4) that such a volume exists if and only if θ is exact. The volume v is said to be *locally invariant under* Ω if for every $p \in F$, there is neighborhood U such that $v(p^*) = v(p)$ provided $\gamma \subset U$ is as in Lemma A. 2. It is clear from (A. 4) that this condition holds if and only if θ is closed. We summarize:

PROPOSITION A. 1. *Let* Ω *be a connection on* B *and let* $v\colon B \to R$ *define a volume on the fibers of* B. *Then there is associated to* Ω *and* v *a* 1-*form* θ *on* F *with the property that:*

(i) *There is a volume invariant under* Ω *if and only if* θ *is exact.*

(ii) *There is a volume locally invariant under* Ω *if and only if* θ *is closed.*

COLUMBIA UNIVERSITY.

REFERENCES.

[1] W. Ambrose and I. Singer, "A theorem on holonomy," *Transactions of the American Mathematical Society*, vol. 5 (1953), pp. 428-443.

[2] A. Denjoy, "Sur les coubres définies par les équations différentelles à la surface du tore," *Journal de Mathématics* (9), vol. 11 (1932), pp. 333-375.

[3] A. Haefliger, "Structures feuilletées et cohomologie à valeur dans un faisceau de groupoides," *Commentarii Mathematici Helvetici*, vol. 32 (1958), pp. 248-329.

[4] ———, "Variétés feuilletées," *Annales della Scuola Normale Superiore di Pisa*, III, XVI (1962), pp. 367-397.

[5] P. Hartman, "On local homeomorphisms of euclidean spaces," *Boletin de la Sociedad Matematica Mexicana* (1960), pp. 220-241.

[6] E. R. van Kampen, "The topological transformations of a simple closed curve into itself," *American Journal of Mathematics*, vol. 57 (1953), pp. 142-152.

[7] E. Lima, "Commuting vector fields on S^3," to appear in the *Annals of Mathematics*.

[8] A. Nijenhuis, "On the holonomy groups of linear connections I, II," *Indagationes Mathematicae*, vol. 16 (1953), pp. 233-250, pp. 241-249.

[9] G. Reeb, "Sur la théorie générale des systèmes dynamiques," *Annales de l'Institut Fourier*, vol. VI (1955), pp. 89-115.

[10] G. Reeb, "Sur les structures feuilletées de co-dimension un et sur un théorem de M. A. Denjoy," *Annales de l'Institut Fourier*, vol. 11 (1961), pp. 185-200.

[11] ———, "Sur certaines propriétés topologiques des variétés feutilletées," *Actualities Scientifiques et Industrielles*, Hermann, Paris (1952).

[12] B. L. Reinhart, "Foliated manifolds with bundle-like metrics," *Annals of Mathematics*, vol. 69 (1959), pp. 119-131.

[13] H. Rosenberg, "The rank of $S^2 \times S^1$," *American Journal of Mathematics*, vol. 87 (1965), pp. 11-24.

[14] R. Sacksteder, "On the existence of exceptional leaves in foliations of co-dimension one," to appear in the *Annales de l'Institut Fourier*, vol. XIV.

[15] ———, "Some properties of foliations," *Annales de l'Institut Fourier*, vol. XIV, pp. 31-35.

[16] R. Sacksteder and A. J. Schwartz, "Limit sets of foliations," to appear in *Annales de l'Institut Fourier*, vol. XV.

[17] A. J. Schwartz, "A generalization of a Poincaré-Bendixson theorem to closed two-dimensional manifolds," *American Journal of Mathematics*, vol. 85 (1963), pp. 453-458.

[18] I. E. Segal, "Invariant measures on locally compact spaces," *Journal of the Indian Mathematical Society*, vol. 13 (1949), pp. 105-130.

[19] C. L. Siegel, "Notes on differential equations on the torus," *Annals of Mathematics*, vol. 46 (1945), pp. 423-428.

[20] S. Sternberg, "Local C^n transformation of the real line," *Duke Mathematical Journal*, vol. 24 (1957), pp. 97-102.

[21] A. Weil, "L'intégration dans le groupes topologiques et ses applications," *Actualites Scientifiques et Industrielles*, Hermann, Paris (1951).

ON THE FIRST COHOMOLOGY OF DISCRETE SUBGROUPS OF SEMI-SIMPLE LIE GROUPS.

By M. S. RAGHUNATHAN.

Introduction. Let G be a connected semi-simple Lie group and Γ a discrete subgroup such that the quotient G/Γ is compact. Let ρ_0 be a finite dimensional representation of G. Our aim in this paper, is to show that for a large class of representations ρ_0, the first cohomology group of Γ with coefficients in the representation ρ_Γ (the restriction of ρ_0 to Γ) is zero (for a precise statement, see Theorem 1). Our results say in particular that if ρ_0 does not contain the trivial representation of G and if no simple component of G is compact or locally isomorphic to $SO_0(n,1)$ or $SU(n,1)$, then this first cohomology group vanishes. Even if G has components locally isomorphic to $SO_0(n,1)$ or $SU(n,1)$ we give a sufficient condition for the vanishing of the cohomology in terms of the highest weights of the irreducible components of the complexification ρ of ρ_0.

The importance of these cohomology groups arises from the role they play in deformation theory; for instance, when ρ_0 is the adjoint representation, these cohomology groups are intimately connected with the theory of deformations of discrete subgroups of Lie groups [6]. It has been proved essentially by A. Weil [8] (see also [5] and [6]) that when ρ_0 is the adjoint representation, this cohomology group vanishes if G has no compact or three dimensional components. This result is a special case of our theorem (see Corollary 1 to Theorem 1). The case of trivial representations has been treated by Y. Matsushima in [4].

1. Statement of the theorem. Throughout this paper G shall denote a connected real semisimple Lie group and Γ a discrete subgroup of G such that G/Γ is compact. We denote by $\mathfrak{g}_0$ the Lie algebra of G.

We denote $SO(n,1)$ (resp. $\mathfrak{so}_0(n,1)$) the subgroup of $GL(n+1,\boldsymbol{R})$ which leaves invariant the quadratic form $x_1^2+x_2^2+\cdots x_n^2-x_{n+1}^2$ in $\boldsymbol{R}^{n+1}$ (resp. the Lie algebra of $SO_0(n,1)$). Similarly $SU(n,1)$ (resp. $\mathfrak{su}(n,1)$) will denote the subgroup of $GL(n+1,\boldsymbol{C})$ leaving invariant the hermitian quadratic form $z_1\bar{z}_1+\cdots+z_n\bar{z}_n-z_{n+1}\bar{z}_{n+1}$ (resp. the Lie algebra of $SU(n,1)$). We denote by τ_0^N the canonical representations of these Lie algebras: $SO_0(n,1)$

Received April 30, 1964.

is a subgroup of $GL(n+1,\boldsymbol{R})$ and $SU(n,1)$ that of $GL(n+1,\boldsymbol{C})$ hence of $GL(2n+2,\boldsymbol{R})$.

In the sequel we denote by $\mathfrak{g}$ the complexification of $\mathfrak{g}_0$ and if ρ_0 is any finite dimensional representation of $\mathfrak{g}_0$ in a (real) vector space V_0, we denote by ρ, the extension of ρ_0 as a representation ρ of $\mathfrak{g}$ in $V=V_0 \underset{R}{\otimes} \boldsymbol{C}$. In the particular case when $\mathfrak{g}_0=\mathfrak{so}_0(n,1)$, τ^N is simply the natural representation of $\mathfrak{g}=\mathfrak{so}(n+1,\boldsymbol{C})$ (the orthogonal Lie algebra in $\boldsymbol{C}^{n+1}$) on $\boldsymbol{C}^{n+1}$. When $\mathfrak{g}_0=\mathfrak{su}(n,1)$, we find that $\mathfrak{g}=\mathfrak{sl}(n+1,\boldsymbol{C})$ and that τ^N breaks up into two components: the natural representation of $\mathfrak{sl}(n+1,\boldsymbol{C})$ in $\boldsymbol{C}^{n+1}$ and its contragredient. We further denote by $\Lambda\rho_0$ any one of the highest weights of ρ. We also set $\Lambda_{\tau N}=\mu_N$ when $\mathfrak{g}_0=\mathfrak{so}_0(n,1)$ or $\mathfrak{su}(n,1)$ and τ^N the representation defined above. With this notation, we can formulate the main result.

THEOREM 1. *Let G be a connected semisimple Lie group and ρ_0 a non-trivial irreducible finite dimensional representation of G in a (real) vector space V_0. Then for any discrete subgroup Γ of G such that G/Γ is compact, $H^1(\Gamma, V_0)=0$ (V_0 being considered as a Γ-module through the restriction ρ_Γ of ρ_0 to Γ) if the following condition is satisfied:*

Let $\mathfrak{g}_0=\sum_i \mathfrak{g}_0{}^i$ be the decomposition of $\mathfrak{g}_0$ into simple components $\mathfrak{g}_0{}^i$ and let $\rho_0{}^i$ denote the restriction to $\mathfrak{g}_0{}^i$ of the representation ρ_0 of $\mathfrak{g}_0$ induced by the representation of G; there should exist one i such that $\mathfrak{g}_0{}^i$ is non-compact, $\rho_0{}^i$ is non-trivial and the pairs $(\mathfrak{g}_0{}^i, \Lambda\rho_0{}^i)$ be different from $(\mathfrak{so}_0(n,1), m\cdot\mu_N)$ and $(\mathfrak{su}(n,1), m\cdot\mu_N)$ (m an integer).

In particular, if G has no component locally isomorphic to $SO_0(n,1)$ or $SU(n,1)$, then $H^1(\Gamma, V_0)=0$.

COROLLARY. *If G has no compact or three dimensional components, then $H^1(\Gamma,\mathfrak{g}_0)=0$, where $\mathfrak{g}_0$, the Lie algebra of G, is considered as a Γ-module through the restriction to Γ of the adjoint representation of G in $\mathfrak{g}_0$.*

Proof of Corollary. It is easy to check that every irreducible component of the adjoint representation (which is completely reducible) satisfies the condition in the theorem except in the case one of the components is three dimensional, i.e., is locally isomorphic to $SU(1,1)$ or what is the same $SO(2,1)$. Hence the corollary (for more details we refer to the end of § 4).

We make a few preliminary reductions. A theorem of Matsushima and Murakami [5] says the following.

PROPOSITION 1. *Let $\mathfrak{g}_0=\mathfrak{k}_0+\mathfrak{p}_0$ be a Cartan-decomposition of $\mathfrak{g}_0$. Let $(X_\alpha)_{1\le\alpha\le r}$ and $(X_i)_{r+1\le i\le N}$ be respectively bases of $\mathfrak{k}_0$ and $\mathfrak{p}_0$ such that*

$B_{\mathfrak{g}_0}(X_\alpha, X_\beta) = -\delta_{\alpha\beta}$ *and* $B_{\mathfrak{g}_0}(X_i, X_j) = \delta_{ij}$ *where* $B_{\mathfrak{g}_0}$ *is the Killing form of* $\mathfrak{g}_0$. *Suppose further that* G *is linear and that* Γ *contains no elements of finite order. Then* Γ *acts without fixed points (on the left) on the symmetric space* $X = G/K$, *and can therefore be identified with the fundamental group of the quotient manifold* $\Gamma\backslash X$. *Let* $\mathbf{L}_{\rho_0}$ *be the locally constant sheaf on* $\Gamma\backslash X$ *associted to the representation* ρ_0. *Then if the operator* $T^1{}_{\rho_0}$ *on the space* $\mathrm{Hom}(\mathfrak{p}_0, V_0)$ *(which is symmetric w.r.t. a suitable scalar product) defined below has all its eigenvalues positive, then* $H^1(\Gamma\backslash X, \mathbf{L}_{\rho_0}) = 0$. *For* $\eta_0 \in \mathrm{Hom}(\mathfrak{p}_0 V_0)$ *and* $Y \in \mathfrak{p}_0$, *we define*

$$T^1{}_{\rho_0}(\eta_0)(Y) = \sum_{k=r+1}^{N} \rho_0(X_k)^2 \cdot \eta_0(Y) + \sum_{k=r+1}^{N} \rho_0([Y, X_k])\eta_0(X_k).$$

(This result is part of Theorem 7.1 of [5]; we have stated the theorem for the special case of 1-forms).

If as in the above proposition, we assume that Γ contains no elements of finite order, i.e., that Γ can be identified with the fundamental group of $\Gamma\backslash X$, it follows from a well known theorem due to S. Eilenberg that $H^1(\Gamma, V_0) \cong H^1(\Gamma\backslash X, \mathbf{L}_{\rho_0})$. We conclude thus

PROPOSITION 2. *If* G *is linear and* Γ *has no elements of finite order and* $T^1{}_{\rho_0}$ *is positive definite (w.r.t. a suitable scalar product), then* $H^1(\Gamma, V_0) = 0$.

Again, by a theorem due to A. Selberg, Γ contains a normal subgroup Γ' of finite index, such that Γ' has no elements of finite order. We conclude then without any difficulty the

PROPOSITION 3. *If* G *is linear and* Γ *is a discrete subgroup such that* G/Γ *is compact and if* $T^1{}_{\rho_0}$ *is positive definite,* $H^1(\Gamma, V_0) = 0$.

We will next get rid of the condition that G is linear. In other words, we have

PROPOSITION 4. G *be any semisimple Lie group and* Γ *a discrete subgroup such that* G/Γ *is compact. Let* ρ_0 *be a non-trivial irreducible finite dimensional representation of* G *in a vector space* V_0. *Then* V_0 *is a* Γ*-module through the restriction* ρ_Γ *of* ρ *to* Γ. *If* $T^1{}_{\rho_0}$ *is positive definite, then* $H^1(\Gamma, V_0) = 0$. *(We assume that* $\rho(G)$ *is non-compact.)*

Proof. Assume that G has no compact components. Let $N(\Gamma)$ be the normaliser of Γ. It is known (see Corollary 1, Appendix II, [8]) that $N(\Gamma)$ is discrete. Moreover, it is clear that the centre of G is contained in $N(\Gamma)$.

Now it is well known that there is a quotient G' of G by a discrete central subgroup, which is linear such that ρ_0 possess down to a representation ρ'_0 of G'. Let $p: G \to G'$ be the covering map. Let $p(\Gamma) = \Gamma'$. Then $p^{-1}(\Gamma') \subset N(\Gamma)$; it follows then that $p^{-1}(\Gamma')$ is discrete and hence so is Γ'. Moreover, clearly G'/Γ' is compact. By Proposition 3, we have $H^1(\Gamma', V_0) = 0$. Now Γ' is a quotient of Γ by a central subgroup, say H. We have then for $x \in \Gamma$, $y \in H$ and a cocycle f on Γ with values in V_0,

$$f(x) + x \cdot f(y) = f(xy) = f(yx) = f(y) + yf(x)$$

and since y is clearly in the kernel of ρ_0, we have

$$x \cdot f(y) = f(y).$$

It follows that $f(y) = 0$. (Theorem 1, Appendix II, [8]: ρ_Γ is irreducible). It is then immediate that

$$f(xy) = f(x) \text{ for } x \in \Gamma, y \in H.$$

Hence f goes down to Γ' as a cocycle on Γ'. Since $H^1(\Gamma', V_0) = 0$, we conclude that f is a coboundary, i.e. every cocycle on Γ is a coboundary. An obvious modification takes care of the case when G has compact components. Hence the proposition.

Proposition 4 reduces Theorem 1 to proving the following

THEOREM 1′. *Let $\mathfrak{g}_0$ be a semisimple Lie algebra and ρ_0 a non-trivial irreducible representation. Let $\mathfrak{g}_0 = \sum_i \mathfrak{g}_0{}^i$ be the decomposition of $\mathfrak{g}_0$ into simple components and let $\rho_0{}^i = \rho_0 |_{\mathfrak{g}_0{}^i}$. Then if there is an i such that $\mathfrak{g}_0{}^i$ is non-compact $\rho_0{}^i$ is non-trivial and none of the pairs $(\mathfrak{g}_0{}^i, \Delta_{\rho_0{}^i})$ are of the form $(\mathfrak{so}_0(n,1), m \cdot \mu_N)$ or $(\mathfrak{su}(n,1), m \cdot \mu_N)$, then $T^1{}_{\rho_0}$ is positive definite.*

We make one more preliminary reduction: we will throw the operator $T^1{}_{\rho_0}$ into a convenient form and then further reduce it to proving the Theorem 1′ in the case when $\mathfrak{g}$ is simple.

Let $(X_i{}^*)_{r+1 \leq i \leq N}$ be the dual basis of $(X_i)_{r+1 \leq i \leq N}$ of $\mathfrak{p}_0$ ($X_i{}^*$ form a basis of the dual $\mathfrak{p}_0{}^*$). Let $\eta_0 = \sum_{r+1}^{N} v_i \otimes X_i{}^*$, $v_i \in V_0$. (We identify $\mathrm{Hom}(\mathfrak{p}_0, V_0)$ with $V_0 \otimes \mathfrak{p}_0{}^*$). We have then

$$\begin{aligned} T_{\rho_0}{}^1(\eta_0)(X_i) &= \{\sum_{r+1}^{N} \rho_0(X_k)^2\} \eta_0(X_i) + \sum_{k=r+1}^{N} \rho_0([X_i, X_k]) \eta_0(X_k) \\ &= \{\sum_{k=r+1}^{N} \rho_0(X_k)^2\}(v_i) + \sum_{k=r+1}^{N} \sum_{\alpha=1}^{r} C_{ik}{}^\alpha \rho_0(X_\alpha)(v_k). \end{aligned}$$

i.e.

$$T_{\rho_0}{}^1(\eta_0) = \{\sum_i \sum_{k=r+1}^{N} \rho_0(X_k)^2(v_i) + \sum_{k=r+1}^{N} \sum_{\alpha=1}^{r} C_{ik}{}^{\alpha}\rho_0(X_\alpha)(v_k)\} \otimes X_i^*$$

$$= \sum_j \{\sum_{k=r+1}^{N} \rho_0(X_k)^2 \otimes 1 + \sum_{\alpha=1}^{r} \rho_0(X_\alpha) \otimes \sigma^*(X_\alpha)\}(v_j \otimes X_j^*)$$

where σ^* denotes dual of the adjoint representation σ of k_0 in p_0 and for $1 \leqq \lambda \leqq N$, we set $[X_\lambda, X_\mu] = \sum C_{\lambda\mu}{}^{\nu} X_\nu$; then denoting by i, j, k etc. (resp. α, β, γ etc.) indices running from $r+1$ to N (resp. 1 to r), we have

$$[X_i, X_j] = \sum_{\alpha=1}^{r} C_{ij}{}^{\alpha} X_\alpha$$

$$[X_\alpha, X_j] = \sum_{k=r+1}^{N} C_{\alpha j}{}^{k} X_k$$

$$[X_\alpha, X_\beta] = \sum_{\gamma=1}^{r} C_{\alpha\beta}{}^{\gamma} X_\gamma$$

with, moreover, the condition $C_{ij}{}^{\alpha} = C_{\alpha j}{}^{i}$; hence we have,

$$\{\sigma^*(X_\alpha)\}(X_i^*) = \sum_j C_{ji}{}^{\alpha} X_j^*.$$

We denote now the element $-\sum_{\alpha=1}^{r} X_\alpha^2$ of $U(\mathfrak{k}_0) \subset U(\mathfrak{g}_0)$, the enveloping algebra of $\mathfrak{g}_0$ by c', and the element $\sum_{r+1}^{N} X_i^2 - \sum_{\alpha=1}^{r} X_\alpha^2$ of $U(\mathfrak{g}_0)$ by c. We have also

$$2\rho_0(X_\alpha) \otimes \sigma^*(X_\alpha) = (\rho_0 \otimes \sigma^*)(X_\alpha^2) - (1 \otimes \sigma^*)(X_\alpha^2) - (\rho_0 \otimes 1)(X_\alpha^2)$$

where all the representations $(\rho_0 \otimes 1)$, $(1 \otimes \sigma^*)$, $(\rho_0 \otimes \sigma^*)$ are assumed extended to the corresponding enveloping algebras and the extensions are denoted by the same symbols. It follows then that

$$2T_{\rho_0}{}^1 = 2(\rho_0 \otimes 1)(c) + (1 \otimes \sigma^*)(c') = (\rho_0 \otimes 1)(c') - (\rho \otimes \sigma^*)(c')$$

We replace in the sequel σ^* by σ: this may be done since the two representations are equivalent. We may therefore write

$$2T_{\rho_0}{}^1 = 2(\rho_0 \otimes 1)(c) + (1 \otimes \sigma)(c') - (\rho_0 \otimes 1)(c') - (\rho_0 \otimes \sigma)(c').$$

Let now $\mathfrak{g}_0 = \sum_t \mathfrak{g}_0{}^t$ be a decomposition of $\mathfrak{g}_0$ into its simple components and for each t, let $\mathfrak{g}_0{}^t = \mathfrak{k}_0{}^t + \mathfrak{p}_0{}^t$ be a Cartan-decomposition of $\mathfrak{g}_0{}^t$. Let $(X_\alpha{}^t)_{1 \leq \alpha \leq r_t}$, $(X_i{}^t)_{r_t+1 \leq i \leq N_t}$ be bases of $\mathfrak{k}_0{}^t$ and $\mathfrak{p}_0{}^t$ respectively such that $B_{\mathfrak{g}_0}(X_\alpha{}^t, X_\beta{}^t) = -\delta_{\alpha\beta}$, $B_{\mathfrak{g}_0}(X_i{}^t, X_j{}^t) = \delta_{ij}$. Let further

$$c_t = \sum_{i=r_t+1}^{N_t} (X_i{}^t)^2 - \sum_{\alpha=1}^{r_t} (X_\alpha{}^t)^2$$

and

$$c_t' = -\sum_{\alpha=1}^{r_t} (X_\alpha^t)^2.$$

We have then

$$2T_{\rho_0}^1 = \sum_t \{2(\rho_0 \otimes 1)(c_t) + (1 \otimes \sigma)(c_t') - (\rho_0 \otimes 1)(c_t') - (\rho_0 \otimes \sigma)(c_t')\}.$$

To prove that $T^1{}_{\rho_0}$ is strictly positive, it is sufficient that each one of the terms under the summation sign is a non-negative symmetric endomorphism of $V_0 \otimes \mathfrak{p}_0$ (w. r. t. a suitable scalar product) and there be one t such that

$$2T_{\rho_t}^1 = 2(\rho_0 \otimes 1)(c_t) + (1 \otimes \sigma)(c_t') - (\rho_0 \otimes 1)(c_t') - (\rho_0 \otimes \sigma)(c_t')$$

is strictly positive. Now if ρ_0 is irreducible, the representation $\rho_t = \rho_0 \,|_{\mathfrak{g}_0^t}$ is a direct sum of isomorphic copies of the same irreducible representation of $\mathfrak{g}_0^t$ (see [1], p. 41). Further $T_{\rho_t}^1 = 0$ if $\mathfrak{g}_0^t$ is compact or if $\rho(\mathfrak{g}_0^t) = 0$. If $\rho(\mathfrak{g}_0^t) \neq 0$, $W = \{v \mid \mathfrak{p}_0^t v = 0\}$ being stable under $\mathfrak{g}_0$ (as is easily seen) is $\{0\}$. Finally for a suitable scalar product $\rho(x_i)^2$ are symmetric and non-negative. Since on $V_0 \otimes \mathfrak{p}_0^s$ $(s \neq t)$, $T\rho_t = \sum_{v_t+1}^{N_t} \rho(x_i)^2 \otimes 1$ $T\rho_t > 0$ on $V_0 \otimes \mathfrak{p}_0^s$ We are thus reduced to proving the following.

THEOREM 1″. *Let $\mathfrak{g}_0$ be a non-compact real simple algebra with a Cartan-decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 \oplus \mathfrak{p}_0$. Let $(X_\alpha)_{1 \leqq \alpha \leqq r}$ and $(X_i)_{r+1 \leqq i \leqq N}$ be bases of k_0 and $\mathfrak{p}_0$ respectively such that $B_{\mathfrak{g}_0}(X_\alpha, X_\beta) = -\delta_{\alpha\beta}$, $1 \leqq \alpha, \beta \leqq r$ and $B_{\mathfrak{g}_0}(X_i, X_j) = \delta_{ij}$, $r+1 \leqq i, j \leqq N$. Let*

$$c = \sum_{i=r+1}^{N} X_i^2 - \sum_{\alpha=1}^{r} X_\alpha^2 \text{ and } c' = -\sum_{\alpha=1}^{r} X_\alpha^2.$$

Then if ρ_0 is any irreducible non-trivial representation the operator

$$T_{\rho_0} = 2(\rho_0 \otimes 1)(c) + (1 \otimes \sigma)(c') - (\rho_0 \otimes 1)(c') - (\rho_0 \otimes \sigma)(c')$$

has all the eigenvalues non-negative and strictly positive if further

(i) *$\mathfrak{g}_0$ is not one of the algebra $so_0(n,1)$ or $su(n,1)$.*

or (ii) *$\mathfrak{g}_0 = \mathfrak{so}(n,1)$ or $\mathfrak{su}(n,1)$; $\Lambda_{\rho_0} \neq m\mu_N$ for any m.*

Remark. Proposition 4 can be proved directly without using the results of A. Weil and A. Selberg quoted; instead of working with the symmetric space as is done by Matsushima and Murakami in [5] we should work directly on the group.

2. Some lemmas on representations of semisimple algebras. Let $\mathfrak{g}$ be semisimple Lie algebra over $\boldsymbol{C}$. Let $B_{\mathfrak{g}}(x,y)$ or $\langle x,y\rangle_{\mathfrak{g}}$ or simply $\langle x,y\rangle$

when there is no ambiguity, be the Killing form on $\mathfrak{g}$. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ and Δ the system of roots of $\mathfrak{g}$ with respect to $\mathfrak{h}$. Let $\mathfrak{g}^\alpha$ denote the space of root vectors of α. For $\alpha \in \Delta$, let $E_\alpha \in \mathfrak{g}^\alpha$ be so chosen that $\langle E_\alpha, E_{-\alpha}\rangle = 1$ and let $[E_\alpha, E_{-\alpha}] = H_\alpha$. Then H_α is the unique element of $\mathfrak{h}$ such that $\langle H, H_\alpha\rangle = \alpha(H)$. We denote by $\mathfrak{h}^*$ the real substance $\sum_{\alpha\in\Delta} RH_\alpha$ of $\mathfrak{h}$; then it is well known that $\mathfrak{h}^*$ has the same dimension over $\boldsymbol{R}$ as $\mathfrak{h}$ over $\boldsymbol{C}$, that every $\alpha \in \Delta$ takes real values on $\mathfrak{h}^*$, and that $B_\mathfrak{g}$ is positive definite on $\mathfrak{h}^*$.

LEMMA 1. *Let $\mathfrak{g} = \sum \mathfrak{g}_i$ be a decomposition of $\mathfrak{g}$ into simple components. Let A be a nondegenerate invariant bilinear form on $\mathfrak{g}$. Then A is a scalar multiple of the Killing form $B_\mathfrak{g}$ on each simple component $\mathfrak{g}_i$. Further, if A is positive definite on $\mathfrak{h}^*$, then these scalars are positive. Also, the $\mathfrak{g}_i$ are mutually orthogonal with respect to the form $A(x,y)$.*

Proof. That $A\,|_{\mathfrak{g}_i}$ is a scalar multiple of $B_{\mathfrak{g}_i}$ is an immediate consequence of Schur's lemma. We have further $B_{\mathfrak{g}_i} = B_\mathfrak{g}\,|_{\mathfrak{g}_i}$ and hence $B_{\mathfrak{g}_i}$ is positive definite on $\mathfrak{h}^* \cap \mathfrak{g}_i$; hence the scalar is positive. If $x, y \in \mathfrak{g}_i$, $z \in \mathfrak{g}_j$, $i \neq j$, we have, since A is invariant, $A([x,y],z) = -A(y,[x,z]) = 0$ and since $[\mathfrak{g}_i, \mathfrak{g}_i] = \mathfrak{g}_i$, $A(\mathfrak{g}_i, \mathfrak{g}_j) = 0$. Hence the last assertion.

LEMMA 2. *Let there be given a lexicographic order on the dual $\mathfrak{h}^{*\prime}$ of $\mathfrak{h}^*$ and A be a nondegenerate invariant bilinear form on $\mathfrak{g}$ which is positive definite on $\mathfrak{h}^*$. Let Δ^+ be the system of positive roots and Λ denote the highest weight of a representation of $\mathfrak{g}$ with respect to this order. Then for $\alpha \in \Delta^+$, $A(\Lambda, \alpha) \geqq 0$.*

Proof. We have denoted by A also the canonically induced bilinear form on $\mathfrak{h}^{*\prime}$. In view of Lemma 1, it is sufficient to prove the lemma for $A = B_\mathfrak{g}$. In this case the lemma is well known. (See [7] Theorem 2, exposé 17).

LEMMA 3. *Let A be a nondegenerate invariant bilinear form on $\mathfrak{g}$. Let $(X_i)_{1\leq i\leq n}$ and $(X_i')_{1\leq i\leq n}$ be two bases of $\mathfrak{g}$ such that $A(X_i, X_j') = \delta_{ij}$; then $c_A = \sum_{i=1}^{n} X_i \cdot X_i'$ is a central element of $U(\mathfrak{g})$, the enveloping algebra of $\mathfrak{g}$ and is independent of the choice of such a pair of bases.*

For a proof we refer to [7, p. 3, exposé 4].

We call c_A the Casimir element corresponding to A and denote $c_{B_\mathfrak{g}}$ by c (for the Lemma, $\mathfrak{g}$ need not be semisimple).

LEMMA 4. *Let ρ be a finite dimensional irreducible representation of $\mathfrak{g}$ and c_A the Casimir element defined above. Then*

$$\rho(c_A) = \{A(\Lambda,\Lambda) + \sum_{\alpha\in\Delta} A(\Lambda,\alpha)\}\cdot Id.,$$

where Λ is the highest weight of ρ (for a lexicographic order on $\mathfrak{h}^*$) and Id. is the Identity operator. (We denote by A, the scalar product induced by A on $\mathfrak{h}^{*\prime}$ also).

Proof. In view of Lemma 1, we may assume that $A = B_{\mathfrak{g}}$. Let then $\{H_i\}_{1 \leq i \leq l}$ be an orthonormal basis of $\mathfrak{h}^*$. Then the bases

$$(\{H_i\}_{1\leq i \leq l}, \{E_\alpha\}_{\alpha \in \Delta}) \text{ and } (\{H_i\}_{1 \leq i \leq l}, \{E_\alpha\}_{-\alpha \in \Delta})$$

satisfy the conditions on $\{X_i\}_{1\leq i\leq l}$, $\{X_i'\}_{1\leq i \leq l}$ of Lemma 3 for $A = B_{\mathfrak{g}}$. Hence $c_{B_{\mathfrak{g}}} = c_B = \sum_{i=1}^{l} H_i^2 + \sum_{\alpha \in \Delta^+} (E_\alpha E_{-\alpha} + E_{-\alpha} E_\alpha)$, Δ^+ designating the system of positive roots. If then v is the highest weight vector, we have $E_\alpha v = 0$ for $\alpha \in \Delta^+$ so that $(E_\alpha E_{-\alpha} + E_{-\alpha}E_\alpha)v = (E_\alpha E_{-\alpha} - E_{-\alpha}E_\alpha)v = [E_\alpha, E_{-\alpha}]v = H_\alpha v$ (we write Xv for $\rho(X)v$) so that

$$\rho(c_{B_{\mathfrak{g}}})v = \{\sum_{i=1}^{l} H_i^2 + \sum_{\alpha \in \Delta^+} H_\alpha\}v = \sum_{i=1}^{l} \Lambda(H_i)^2 v + \sum_{\alpha \in \Delta^+} \Lambda(H_\alpha) v$$
$$= \{\langle \Lambda, \Lambda \rangle + \sum_{\alpha \in \Delta^+} \langle \Lambda, \alpha \rangle\} v$$

by the choice of the H_α. This proves the lemma. ($\rho(c_B)$ is a scalar operator).

LEMMA 5. *Let ρ_1 and ρ_2 be irreducible representations of $\mathfrak{g}$ in finite dimensional vector spaces with highest weights Λ_1, Λ_2. Suppose $\Lambda_1 - \Lambda_2 = \sum_{i=1}^{l} m_i \alpha_i$ where α_i are simple roots of $\mathfrak{g}$ and $m_i \geqq 0$. Then if A is an invariant nondegenerate form on $\mathfrak{g}$ such that $A\,|_{\mathfrak{h}^*}$ is positive definite and λ_i, $i = 1, 2$, are defined by $\lambda_i \cdot$ (Identity) $= \rho_i(c_A)$, we have $\lambda_1 \geqq \lambda_2$; if $\Lambda_1 \neq \Lambda_2$, $\lambda_1 > \lambda_2$.*

Proof. Once again, by Lemma 1, we may assume $A = B_{\mathfrak{g}}$ and the have in this case

$$\rho_1(c) = \rho_1(c_{B_{\mathfrak{g}}}) = \langle \Lambda_1, \Lambda_1 \rangle + \sum_{\alpha \in \Delta^+} \Lambda_1(H_\alpha) \text{ and}$$

$$\rho_2(c) = \rho_2(c_{B_{\mathfrak{g}}}) = \langle \Lambda_2, \Lambda_2 \rangle + \sum_{\alpha \in \Delta^+} \Lambda_2(H_\alpha).$$

We have $\Lambda_1 = \Lambda_2 + \sum_{i=1}^{l} m_i \alpha_i$, so that,

$$\langle \Lambda_1, \Lambda_1 \rangle = \langle \Lambda_2, \Lambda_2 \rangle + 2\langle \Lambda_2, \sum_{i=1}^{l} m_i \alpha_i \rangle + \langle \sum m_i \alpha_i, \sum m_i \alpha_i \rangle.$$

Now $\langle \Lambda_2, \alpha_i \rangle \geqq 0$ for every i since Λ_2 is the highest weight of a finite dimensional representation (Lemma 2). Further if $\sum_{\alpha \in \Delta^+} H_\alpha$ is denoted H_ρ, then

$\alpha(H_\rho) > 0$ for every $\alpha \in \Delta^+$ so that $(\Lambda_2 - \Lambda_1)(H_\rho) \leqq 0$ (see [7], Lemma 2, exposé No. 19). Hence the lemma.

LEMMA 6. *Let ρ be an irreducible faithful finite dimensional representation of $\mathfrak{g}$ and Λ its highest weight. Let $\alpha_1, \cdots, \alpha_l$ be the system of simple roots. Let*

$$\Lambda = \sum_{i=1}^{l} m_i \alpha_i.$$

We assert that $m_i > 0$ for all i.

Proof. We first prove that all the $m_i \geqq 0$: if possible let

$$\Lambda = \sum_{i \in I_1} m_i \alpha_i - \sum_{j \in I_2} m_j \alpha_j$$

with $m_j, m_i \geqq 0$ and $I_1 \cap I_2 = \emptyset$ and for at least one $j \in I_2$, $m_j \neq 0$; then we have (since Λ is the highest weight),

$$0 \leqq \langle \Lambda, \sum_{j \in I_2} m_j \alpha_j \rangle = \langle \sum_{i \in I_1} m_i \alpha_i, \sum_{j \in I_2} m_j \alpha_j \rangle - \langle \sum_{j \in I_2} m_j \alpha_j, \sum_{j \in I_2} m_j \alpha_j \rangle$$

and since $\langle \alpha_i, \alpha_j \rangle \leqq 0$ for $i \in I_1$, $j \in I_2$, the simple roots being distinct, the right side is strictly negative ($m_j \neq 0$ for at least one $j \in I_2$), a contradiction. Hence

$$\Lambda = \sum_{i=1}^{l} m_i \alpha_i, \qquad m_i \geqq 0.$$

We next reduce the proof to the case of simple $\mathfrak{g}$. Suppose $\mathfrak{g} = \sum_i \mathfrak{g}^i$, $\mathfrak{g}^i$ simple components of g, then $\mathfrak{h} = \sum_i \mathfrak{h}^i$ where $\mathfrak{h}^i = \mathfrak{h} \cap \mathfrak{g}^i$. If Λ is zero on one of the $\mathfrak{h}^i$ say $\mathfrak{h}^{i_0}$ then $\rho \mid \mathfrak{g}_{i_0}$ is trivial contradicting the faithfulness of ρ. Hence $\Lambda \mid_{\mathfrak{h}^i}$ is non-zero and is the highest weight of a representation; since the system of simple roots of $\mathfrak{g}$ is the disjoint union of the systems of simple roots of the $\mathfrak{g}^i$, we are reduced to the case of a nontrivial highest weight of a simple algebra.

Assuming then that $\mathfrak{g}$ is simple, if possible, let $\Lambda = \sum_{i \in I_1} m_i \alpha_i$, $m_i > 0$, where I_1 is not the entire system of simple roots. Since $\mathfrak{g}$ is simple, there exists $i_1 \in I_1$, $j_1 \notin I_1$ such that $\langle \alpha_{i_1}, \alpha_{j_1} \rangle \neq 0$. But then

$$0 \leqq \langle \Lambda, \alpha_{j_1} \rangle = \sum_{i \in I_1} m_i \langle \alpha_i, \alpha_{j_1} \rangle < 0$$

(the right side being < 0 because, $\langle \alpha_i, \alpha_{j_1} \rangle \leqq 0$ for all $i \neq j$, and $\langle \alpha_{i_1}, \alpha_{j_1} \rangle < 0$ and for $i \in I_1$, $m_i > 0$), a contradiction. This proves the lemma.

COROLLARY. *Let ρ_1 and ρ_2 be non-trivial representations of $\mathfrak{g}$ such that ρ_1 is a faithful. Let Λ_1, Λ_2 be the highest weights of ρ_1 and ρ_2 respectively. Then for any invariant scalar product A on $\mathfrak{g}$ which is positive definite on $\mathfrak{h}^*$,*

$$A(\Lambda_1, \Lambda_2) > 0.$$

Proof. As before we may assume that $A = B\mathfrak{g}$. Let $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2$ where $\mathfrak{g}_1$ is the kernel of ρ_2. Let $\alpha_1, \cdots, \alpha_{l_1}$ be the simple roots of $\mathfrak{g}_1$ and $\beta_1, \cdots, \beta_{l_2}$ those of $\mathfrak{g}_2$. Then we have

$$\Lambda_1 = \sum_{i=1}^{l_1} m_i \alpha_i + \sum_{j=1}^{l_2} n_j \beta_j$$

$$\Lambda_2 = \sum_{k=1}^{l_2} q_k \beta_k$$

where $m_i, n_j, q_k > 0$. Hence since $\langle \Lambda_1, \beta_j \rangle_{\mathfrak{g}} > 0$ for at least one j, and $\langle \Lambda_1, \beta_j \rangle_{\mathfrak{g}} \geqq 0$ for all j,

$$\langle \Lambda_1, \Lambda_2 \rangle > 0.$$

Hence the corollary.

We adopt the following notation. For a linear form λ on $\mathfrak{h}$, we define $f_i(\lambda)$ by setting

$$\lambda = \sum_{i=1}^{l} f_i(\lambda) \alpha_i$$

where $\alpha_1, \cdots, \alpha_l$ are the simple roots of $\mathfrak{g}$. With this notation we have

LEMMA 7. *Let $\mathfrak{g}$ be simple and ρ an irreducible representation with highest weight Λ. Suppose now that for every i with $\langle \Lambda, \alpha_i \rangle \neq 0$ (hence > 0; Lemma 2), $f_i(\Lambda) \leqq 1$, equality holding for at most one such i, then $(\mathfrak{g}, \rho)$ must be one of the list below.*

A. *Rank* $\mathfrak{g} = 1$; $\mathfrak{g} = \mathfrak{sl}(2, \boldsymbol{C})$ *and ρ is either the natural representation or the adjoint representation.*

B. *Rank* $\mathfrak{g} = 2$; $\mathfrak{g} = \mathfrak{so}(5, \boldsymbol{C})$ *and ρ the natural representation,*
or $\mathfrak{g} = \mathfrak{sp}(2, \boldsymbol{C})$ *and ρ the natural representation,*
or $\mathfrak{g} = \mathfrak{sl}(3, \boldsymbol{C})$ *and ρ the natural representation or its dual.*

C. *Rank* $\mathfrak{g} = 3$; $\mathfrak{g} = \mathfrak{so}(2n, \boldsymbol{C})$ *and ρ the natural representation,*
or $\mathfrak{g} = \mathfrak{so}(2n+1, \boldsymbol{C})$ *and ρ the natural representation,*
or $\mathfrak{g} = \mathfrak{sp}(n, \boldsymbol{C})$ *and ρ the natural representation,*
or $\mathfrak{g} = \mathfrak{sl}(n, \boldsymbol{C})$ *and ρ is the natural representation or its dual.**

* In the other cases ρ is self-dual.

Proof. The proof is obvious for rank $\mathfrak{g}=1$. For the proof when rank $\mathfrak{g}>1$, we use the following lemma.

LEMMA 8. *Let $S=\{\alpha_1\cdots,\alpha_k\}$ be a subset of the system of simple roots of a simple algebra $\mathfrak{g}$. We assume that S is "connected," i.e., for each integer i with $1\leqq i\leqq k$, there are integers $i_1,\cdots,i_p$, $1\leqq i_r\leqq k$ such that $i_1=1$, $i_p=i$ and $\langle\alpha_{i_r},\alpha_{i_{r+1}}\rangle\neq 0$ (hence <0). Suppose now that Λ is a linear form on $\mathfrak{h}$ such that*

(i) $$\Lambda=\sum_{i=1}^{k} m_i\alpha_i,\ m_i>0,\ m_1\leqq 1$$

(ii) $2\langle\Lambda,\alpha_i\rangle/\langle\alpha_i,\alpha_i\rangle$ *is zero for $i\neq 1$ and a positive integer for $i=1$.*

Then we have only the following possibilities.

($\mathbf{P}_1$) $k=1$; $\Lambda=\alpha_1$ or $\frac{1}{2}\alpha_1$.

($\mathbf{P}_2$) $k=2$; we have either

(i) $\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle$ and $\Lambda=\alpha_1+\alpha_2$,

or **(ii)** $2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle$ and $\Lambda=\alpha_1+\frac{1}{2}\alpha_2$,

or **(iii)** $\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle$ and $\Lambda=\frac{2}{3}\alpha_1+\frac{1}{3}\alpha_2$.

($\mathbf{P}_3$) $k\geqq 3$; after a reordering $\beta_1,\cdots,\beta_k$ of the roots $\alpha_1,\cdots,\alpha_k$ with $\alpha_1=\beta_1$, we have either

(i) $\langle\beta_i,\beta_j\rangle=0$ except in the following cases: we assume $i<j$; $j=i+1\leqq k-1$; $i=k-2$, $j=k-1$; $i=k-2$, $j=k$. Also, we have

$$\langle\beta_i,\beta_i\rangle=\langle\beta_j,\beta_j\rangle \text{ and } \Lambda=\sum_{i=1}^{k-2}\alpha_i+\tfrac{1}{2}(\alpha_{k-1}+\alpha_k)$$

or **(ii)** $\langle\beta_i,\beta_j\rangle=0$ except in the following cases: we assume $i<j$; $j=i+1\leqq k$. Also, $\langle\beta_i,\beta_i\rangle=\langle\beta_j,\beta_j\rangle$ for $i,j<k$

$$\langle\beta_k,\beta_k\rangle=\tfrac{1}{2}\langle\beta_i,\beta_i\rangle \quad i\neq k \text{ and } \Lambda=\sum_{i=1}^{k}\beta_i$$

or **(iii)** $\langle\beta_i,\beta_j\rangle=0$ except when (we assume $i<j$) $j=i+1\leqq k$; we have $\langle\beta_i,\beta_i\rangle=\langle\beta_j,\beta_j\rangle$ for $i,j<k$ and

$$\langle\beta_k,\beta_k\rangle=2\langle\beta_i,\beta_i\rangle \text{ for } i\neq k \text{ and } \Lambda=\sum_{i=1}^{k-1}\alpha_i+\tfrac{1}{2}\alpha_k.$$

or **(iv)** $\langle\beta_i,\beta_j\rangle=0$ (we assume $i<j$) except if $j=i+1\leqq k$; also $\langle\beta_i,\beta_i\rangle=\langle\beta_j,\beta_j\rangle$ for any pair i,j and $\Lambda=\sum_{i=1}^{k}\frac{k-i}{k+1}\cdot\beta_i$.

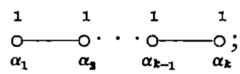
Proof of Lemma 7. Let $\alpha_1,\cdots,\alpha_l$ be the system of simple roots of $\mathfrak{g}$. We define Λ_i by setting $2\langle\Lambda_i,\alpha_j\rangle/\langle\alpha_j,\alpha_j\rangle=\delta_{ij}$; then Λ_i is the highest weight of a representation ρ_i of $\mathfrak{g}$. If now Λ is the highest weight of any representation ρ, we have

$$\Lambda=\sum_{i=1}^{l} m_i\Lambda_i$$

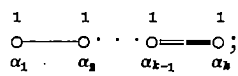
where $m_i=2\langle\Lambda,\alpha_i\rangle/\langle\alpha_i,\alpha_i\rangle$ is a non-negative integer. Let us assume α_1 to be so chosen that (i) $\langle\Lambda,\alpha_1\rangle\neq 0$ and (ii) if $\langle\Lambda,\alpha_i\rangle\neq 0$ for some i, then $f_i(\Lambda)\leqq f_1(\Lambda)$. We will now show that under the assumptions, $\Lambda=\Lambda_1$. Since the m_i are nonnegative integers and $m_1\neq 0$, we see that

$$\text{(i)}\quad 2\langle\Lambda_1,\alpha_i\rangle/\langle\alpha_i,\alpha_i\rangle=\delta_{1i}\quad\text{and}\quad\text{(ii)}\quad f_1(\Lambda_1)\leqq f_1(\Lambda)\leqq 1.$$

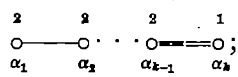
Now, by Lemma 8, we see that only the following things can happen.

a) diagram of $\mathfrak{g}$ is 

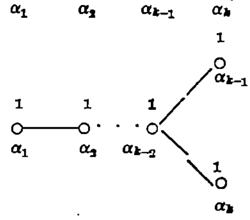
; $\Lambda_1=\sum_{i=1}^{k}\frac{(k+1)}{(k-i)}\alpha_i$.

b) diagram of $\mathfrak{g}$ is 

; $\Lambda_1=\sum_{i=1}^{k-1}\alpha_i+(\alpha_k/2)$.

c) diagram of $\mathfrak{g}$ is 

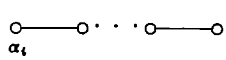
; $\Lambda_1=\sum_{i=1}^{k}\alpha_i$.

d) diagram of $\mathfrak{g}$ is 

; $\Lambda_1=\sum_{k-1}^{i=1}\alpha_i+\alpha_{k-1}/2+\alpha_k/2$.

(To $\mathbf{P}_1$ correspond (a) with $k=1$ and (c) with $k=1$; To $\mathbf{P}_2(\mathbf{i})$ corresponds (c) with $k=2$; To $\mathbf{P}_2(\mathbf{ii})$, corresponds (b) with $k=2$ and to $\mathbf{P}_2(\mathbf{iii})$ (a) with $k=2$. When rank $\mathfrak{g}>3$, $\mathbf{P}_3(\mathbf{i})$, $\mathbf{P}_3(\mathbf{ii})$, $\mathbf{P}_3(\mathbf{iii})$ and $\mathbf{P}_3(\mathbf{iv})$ correspond respectively to (d), (c), (b) and (a) above) (We have, in the above, also assumed the α_i so reordered as to fit into the diagram as above,)),

In the cases (b), (c) and (d), we have $f_1(\Lambda_1)=1$. In view of Lemma 6, $f_1(\Lambda_i)>0$ for every i. Hence since $f_1(\Lambda)=\sum_{i=1}^{k}m_if_1(\Lambda_i)\leqq 1$, we conclude that $m_i=0$ if $i\neq 1$ and $m_1=1$. Hence $\Lambda=\Lambda_1$. In the case (a), $f_1(\Lambda_1)=k/(k+1)$. Now suppose for some i with $1<i\leqq k$, we have $f_i(\Lambda_i)\leqq 1$, then by Lemma 8, by a reordering of $\alpha_1,\cdots\alpha_k$ the diagram of $\mathfrak{g}$ would become



In other words $i = k$. Moreover from Lemma 8, $\Lambda_k = \sum_{i=1}^{k} \frac{i}{k+1} \alpha_i$. It follows that

$$\Lambda = m_1\Lambda_1 + m_k\Lambda_k.$$

Assume that $m_k > 0$ (m_k is an integer.) Then

$$f_1(\Lambda) = m_1 \cdot (k-1)/k + m_k \cdot 1/k \geqq 1;$$

similarly $f_k(\Lambda) \geqq 1$, a contradiction. Hence $m_k = 0$ i. e. $\Lambda = m_1\Lambda_1$.

By a direct comparison of the highest weights we find that to (a), (b), (c) and (d) correspond respectively the following algebras and representations.

(a) $\mathfrak{g} \cong \mathfrak{sl}(k+1, \boldsymbol{C})$; the isomorphism can be given in two natural ways and correspondingly ρ (which has Λ for highest weight) is either the natural representation or its contragredient.

(b) $\mathfrak{g} \cong \mathfrak{sp}(k, \boldsymbol{C})$; through the natural identification ρ can be identified with the natural representation of $\mathfrak{sp}(k, \boldsymbol{C})$ in $\boldsymbol{C}^{2k}$.

(c) $\mathfrak{g} \cong \mathfrak{so}(2k+1, \boldsymbol{C})$ and ρ is then the natural representation in $\boldsymbol{C}^{2k+1}$.

(d) $\mathfrak{g} \cong \mathfrak{so}(2k, \boldsymbol{C})$ and ρ is the natural representation in $\boldsymbol{C}^{2k}$. (Here there are two "natural" isomorphisms of $\mathfrak{g}$ on $\mathfrak{so}(2k, \boldsymbol{C})$ but the representations of $\mathfrak{so}(2k, \boldsymbol{C})$ got through either identification is the natural one).

Proof of Lemma 8. We argue by induction on k. The start of the induction (at $k=3$) is treated at the end. We will over and over again use the following fact.

Remark I. *$2\langle\alpha_i, \alpha_j\rangle/\langle\alpha_j, \alpha_j\rangle$ is a non-positive integer if $i \neq j$.*

Assertion I. If $\langle\alpha_i, \alpha_j\rangle \neq 0$ (hence <0) for some $j \neq i$, we have $2m_i \geqq m_j$; strict inequality holds, if moreover, there is a j_1 with $j_1 \neq i$ or j and $\langle\alpha_i, \alpha_{j_1}\rangle \neq 0$.

Proof. In fact from our assumptions we have using Remark I

$$0 \leqq 2\langle\Lambda, \alpha_i\rangle/\langle\alpha_i, \alpha_i\rangle = 2m_i + 2m_j\langle\alpha_i, \alpha_j\rangle/\langle\alpha_i, \alpha_i\rangle + \lambda$$

where $\lambda \leqq 0$ and <0 if there is a j_1 as in the second assumption of the lemma. From the integrality of $2\langle\alpha_i, \alpha_j\rangle/\langle\alpha_i, \alpha_i\rangle$, we have $2m_i \geqq m_j$, strict inequality holding if there is a j_1 as in the statement of the assertion. (Note. This assertion holds even if $k \leqq 3$.)

Assertion II. Assume $k > 3$; then there is one and only one integer i_0, $1 < i_0 \leqq k$ such that $\langle\alpha_{i_0}, \alpha_1\rangle \neq 0$; also, there exists j_0 with $j_0 \neq i_0$ or 1 and $\langle\alpha_{i_0}, \alpha_{j_0}\rangle \neq 0$.

Proof. If possible, let α_{i_0}, α_j be two roots with $\langle\alpha_{i_0}, \alpha_1\rangle \neq 0$ and $\langle\alpha_j, \alpha_1\rangle \neq 0$, $i_0 \neq 1$, $j \neq 1$, $i_0 \neq j$. We have then (using Remark I)

$$1 \leqq 2\langle\Lambda, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle = 2m_1 + m_{i_0}\cdot 2\langle\alpha_{i_0}, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle + m_j 2\langle\alpha_j, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle + x$$

where $x \leqq 0$; it follows that $2m_1 \geqq m_{i_0} + m_j + 1 - x$. From Assertion I, we have $m_{i_0} \geqq \frac{1}{2}m_1$ and $m_j \geqq \frac{1}{2}m_1$. Since $k > 3$ and S is "connected," there exists an integer i' such that one of α_{i_0} or α_j, say α_{i_0}, is connected to $\alpha_{i'}$. Then, by Assertion I, we have $m_{i_0} > m_1/2$ so that $m_1 > 1 - x \geqq 1$, a contradiction. Hence the assertion.

Assertion III. If the conclusions of Assertion II hold (in particular if $k > 3$) then $m_{i_0} \leqq m_1$ and $2\langle\alpha_{i_0}, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle = -1$.

Proof. From Assertions I and II, we have $m_{i_0} > m_1/2$. Also from Assertion II, we have

$$1 \leqq 2\langle\Lambda, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle = 2m_1 + m_{i_0}\cdot 2\langle\alpha_1, \alpha_{i_0}\rangle/\langle\alpha_1, \alpha_1\rangle$$

Hence if $-(2\langle\alpha_1, \alpha_{i_0}\rangle/\langle\alpha_1, \alpha_1\rangle) \geqq 2$, then

$$1 < 2m_1 - 2\cdot m_1/2 = m_1 \leqq 1,$$

a contradiction. Since $-2\langle\alpha_1, \alpha_{i_0}\rangle/\langle\alpha_1, \alpha_1\rangle$ is a positive integer, it must be 1. Moreover, we have then

$$2m_1 - m_{i_0} = 1.$$

If $m_{i_0} > m_1$, we get $2m_1 - m_1 > 1$ which contradicts the hypothesis. Hence $m_{i_0} \leqq m_1$. This completes the proof.

Assertion IV. $(2\langle\alpha_{i_0}, \alpha_1\rangle/\langle\alpha_{i_0}, \alpha_{i_0}\rangle) = -1$ if $k > 3$ or if the conclusions of Assertion II hold.

Proof. We have

$$0 \leqq 2\langle\Lambda, \alpha_{i_0}\rangle/\langle\alpha_{i_0}, \alpha_{i_0}\rangle = 2m_{i_0} + m_1\cdot 2(\langle\alpha_1, \alpha_{i_0}\rangle/\langle\alpha_{i_0}, \alpha_i\rangle) + x$$

where $x < 0$. (Assertion II combined with Assertion I). Hence if

$$-2\langle\alpha_1, \alpha_{i_0}\rangle/\langle\alpha_{i_0}, \alpha_{i_0}\rangle \geqq 2,$$

we have,

$$2m_{i_0} - 2m_1 > 0, \text{ hence, } 2m_1 - 2m_1 > 0 \text{ (Assertion III)},$$

a contradiction. Hence the assertion.

Consider now, the linear form $\Lambda' = 1/m_1(\Lambda - m_1\alpha_1)$. We assert that

Λ' satisfies the hypotheses of Lemma 8 w.r.t. the connected system of roots $(\alpha_2, \cdots, \alpha_k)$, the role of α_1 being played now by α_{i_0}. In fact

$$2\langle\Lambda', \alpha_{i_0}\rangle/\langle\alpha_{i_0}, \alpha_{i_0}\rangle = \{(2/m_1)\langle\Lambda, \alpha_{i_0}\rangle/\langle\alpha_{i_0}, \alpha_{i_0}\rangle\} - \{2\langle\alpha_1, \alpha_{i_0}\rangle/\langle\alpha_{i_0}, \alpha_{i_0}\rangle\} = 1$$

(we have used Assertions IV and the hypothesis on Λ). Moreover, if $i \neq i_0$ or 1, we have $\langle\Lambda', \alpha_i\rangle = 0$ since $\langle\Lambda, \alpha_i\rangle = 0$ (hypothesis) and $\langle\alpha_1, \alpha_i\rangle = 0$ (Assertion II). Finally

$$\Lambda' = (m_{i_0}/m_1)\alpha_{i_0} + \sum_{i \neq i_0 \text{ or } 1} (m_i/m_1)\alpha_i.$$

and by Assertion III, $(m_{i_0}/m\ ) \leqq 1$.

By induction hypothesis, the Lemma holds for Λ'. Combining then this fact and Assertions III and IV, we find that if Λ' falls in $\mathbf{P}_3(\mathbf{i})$, $\mathbf{P}_3(\mathbf{ii})$, $\mathbf{P}_3(\mathbf{iii})$, or $\mathbf{P}_3(\mathbf{iv})$ then Λ falls in the corresponding case.

We have to start the induction at $k = 3$. Here there are first of all two possibilities.

(a) $\langle\alpha_1, \alpha_2\rangle \neq 0$ while $\langle\alpha_1, \alpha_3\rangle = 0$ (after a suitable reordering if necessary of α_2, α_3).

(b) $\langle\alpha_1, \alpha_2\rangle \neq 0$, $\langle\alpha_1, \alpha_3\rangle \neq 0$. (The third scalar product is zero, since there cannot be a "cycle" in a system of simple roots.)

Assume (a). Since Assertions III, IV hold if the conclusions of Assertion II hold—and this is precisely condition (a)—in this case we reduce the problem to the case of $k = 2$.

Assume (b). We have, then,

$$1 \leqq 2\langle\Lambda, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle = 2m_1 + m_2 \cdot 2\langle\alpha_2, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle + m_3 \cdot 2\langle\alpha_3, \alpha_1\rangle/\langle\alpha_1 \alpha_1\rangle.$$

Since $m_2 \geqq \frac{1}{2}m_1$, $m_3 \geqq \frac{1}{2}m_1$ (Assertion 1), and $m_1 \leqq 1$, we have

$$2\langle\alpha_2, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle = 2\langle\alpha_3, \alpha_1\rangle/\langle\alpha_1, \alpha_1\rangle = -1,$$

and $m_1 = 1$, $m_2 = m_3 = \frac{1}{2}$. Since $\langle\alpha_3, \alpha_2\rangle = 0$, from $m_2 = m_3 = \frac{1}{2}m_1 = \frac{1}{2}$, using $\langle\Lambda, \alpha_2\rangle = 0$, we conclude that

$$2\langle\alpha_3, \alpha_1\rangle/\langle\alpha_3, \alpha_3\rangle = 2\langle\alpha_2, \alpha_1\rangle/\langle\alpha_2, \alpha_2\rangle = -1,$$

i.e., we are in $\mathbf{P}_3(\mathbf{i})$.

We are now left with the case when $k = 2$. We have then $\Lambda = m_1\alpha_1 + m_2\alpha_2$. Now it is well known that $2\langle\alpha_1, \alpha_2\rangle/\langle\alpha_1, \alpha_1\rangle = -1$ or -2 or -3 and also $(2\langle\alpha_1, \alpha_2\rangle/\langle\alpha_1, \alpha_1\rangle)(2\langle\alpha_1, \alpha_2\rangle/\langle\alpha_2, \alpha_2\rangle) = 1, 2,$ or 3.

If now $2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle=-3$ and consequently $2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=-1$, or the reverse happens, then we have either $2m_1-3m_2=1$, $2m_2-m_1=0$ leading to $\frac{1}{2}m=1$ a contradiction; or, $2m_1-m_2=1$, $2m_2-3m_1=0$ leading again to $m_1=2$, a contradiction.

Hence $-2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle$ and $-2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle$ are both integers less than or equal to 2.

a(i). We assume $\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle=-1$. We then have $2m_1-m_2=1$; $2m_2-2m_1=0$; i.e., $m_1=m_2=1$; this is $\mathbf{P}_2(\mathbf{i})$.

a(ii). Assume $2\langle\alpha_1,\alpha\ \rangle/\langle\alpha_2,\alpha_2\rangle=\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle=-1$. In this case, $2m_1-2m_2=1$; $2m_2-m_1=0$; i.e., $m_1=1$, $m_2=\frac{1}{2}$. This is $\mathbf{P}_2(\mathbf{ii})$.

a(iii). $2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=2\langle\alpha_1,\alpha_2\rangle/\langle\alpha_1,\alpha_1\rangle=-1$. We have

$$1=2\langle\Lambda,\alpha_1\rangle/\langle\alpha_1,\alpha_1\rangle=2m_1-m_2 \text{ and}$$
$$0=2\langle\Lambda,\alpha_2\rangle/\langle\alpha_2,\alpha_2\rangle=2m_2-m_1.$$

Hence $m_1=\frac{2}{3}$, $m_2=\frac{1}{3}$. This is the case $\mathbf{P}_2(\mathbf{iii})$.

(Adopting the notation of the earlier part of the proof one concludes that if $k=3$ and if we are in case a, then Λ is in $\mathbf{P}_3(\mathbf{ii})$, $\mathbf{P}_3(\mathbf{iii})$ or $\mathbf{P}_3(\mathbf{iv})$ according as Λ'_0 is in $\mathbf{P}_2(\mathbf{i})$, $\mathbf{P}_2(\mathbf{ii})$ or $\mathbf{P}_2(\mathbf{iii})$.)

LEMMA 9. *Let $\mathfrak{g}$ be semisimple and ρ a faithful irreducible representation. Suppose for some i, $1\leqq i\leqq l$, we have, $\langle\Lambda,\alpha_i\rangle\neq 0$ then $f_i(\Lambda)\geqq\frac{1}{2}$; equality can occur only if $\mathfrak{g}=\mathfrak{sl}(2,\boldsymbol{C})\times\mathfrak{g}_1$ where $\mathfrak{g}_1$ is some semi-simple algebra and α_i is the root corresponding to the component $\mathfrak{sl}(2,\boldsymbol{C})$ and $\rho=\rho_1\otimes\rho_2$ where ρ_1 is the natural representation if $\mathfrak{sl}(2,\boldsymbol{C})$ and ρ_2 is a representation of $\mathfrak{g}_1$.*

Proof. $2\langle\Lambda,\alpha_i\rangle/\langle\alpha_i,\alpha_i\rangle$ is a non-negative integer. Hence,

$$1\leqq 2\langle\Lambda,\alpha_i\rangle/\langle\alpha_i,\alpha_i\rangle=\sum_j f_j(\Lambda)\cdot 2\langle\alpha_i,\alpha_j\rangle/\langle\alpha_j,\alpha_j\rangle$$

The right side is clearly less than or equal to $2f_i(\Lambda)$, i.e., $2f_i(\Lambda)\geqq 1$; moreover, equality can occur if and only if $\langle\alpha_i,\alpha_j\rangle=0$ for $i\neq j$, i.e., α_i is not connected to any other root. The lemma follows quite easily from this.

LEMMA 10. *Let ρ be a faithful irreducible (finite dimensional) representation of a semisimple algebra $\mathfrak{g}$. Let Λ be the highest weight of ρ and let*

$$S_\rho=\{\lambda\mid\lambda \text{ a weight of};\ f_i(\lambda)\geqq 0 \text{ for } 1\leqq i\leqq l\}.$$

(Then $\Lambda\in S_\rho$; see *Lemma* 6). *Let $\mu_\rho=\sum_{\lambda\in S_\rho-\{\Lambda\}}\lambda$. Then $f_i(\mu_\rho)\geqq 0$ for every*

i and $f_i(\mu_\rho) > 0$ for $1 \leqq i \leqq l$ except in the following case: $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2 \times \mathfrak{g}_3$ *where $\mathfrak{g}_1$, $\mathfrak{g}_2$ are direct products of algebras of the form $\mathfrak{sl}(n, \boldsymbol{C})$ and $\mathfrak{g}_3$ is one of the algebras $\mathfrak{sl}(n, \boldsymbol{C})$, $\mathfrak{so}(n, \boldsymbol{C})$ $(n \neq 4)$ or $\mathfrak{sp}(n, \boldsymbol{C})$; further $\rho = \rho_1 \otimes \rho_2 \otimes \rho_3$ where ρ_1 is a tensor product of the natural representations of the components of $\mathfrak{g}_1$, ρ_2 a tensor product of duals of the natural representations of the components of $\mathfrak{g}_0$ and ρ_3 is the natural representation (or its dual if $\mathfrak{g}_3 = \mathfrak{sl}(n, \boldsymbol{C})$) of $\mathfrak{g}_3$.*

Proof. Case i. For some i with $1 \leqq i \leqq 1$, $\langle \Lambda, \alpha_i \rangle \neq 0$ and $f_i(\Lambda) > 1$.

Since $\langle \Lambda, \alpha_i \rangle > 0$ the entire α_i-series,

$$\Lambda, \Lambda - \alpha_i, \cdots, \Lambda - (2\langle \Lambda, \alpha_i \rangle / \langle \alpha_i, \alpha_i \rangle) \cdot \alpha_i$$

are weights of ρ, in particular $\Lambda - \alpha_i$ is. Also, clearly $f_j(\Lambda - \alpha_i) = f_j(\Lambda) > 0$ if $j \neq i$ (see Lemma 6) and $f_i(\Lambda - \alpha_i) = f_i(\Lambda) - 1 > 0$ since $f_i(\Lambda) > 1$. Hence $\Lambda - \alpha_i \in S_\rho$ and $f_j(\mu_\rho) > 0$ for $1 \leqq i \leqq l$.

Case ii. For a pair (i, j), $i \neq j$, $1 \leqq i, j \leqq 1$, $j \leqq 1$, $f_j(\Lambda) = 1$, $f_i(\Lambda) = 1$ and $\langle \Lambda, \alpha_i \rangle \neq 0$, $\langle \Lambda, \alpha_j \rangle \neq 0$.

In this case again consider the "α_i-series," Λ, $(\Lambda - \alpha_i), \cdots, \Lambda - \dfrac{2\langle \Lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \alpha_i$; the entire set are weights of ρ. Also, we have for $\Lambda - \alpha_i$, $f_k(\Lambda - \alpha_i) = f_k(\Lambda)$ for $i \neq k$ while $f_i(\Lambda - \alpha_i) = f_i(\Lambda) - f_i(\alpha_i) = 0$. Hence $\Lambda - \alpha_i \in S_\rho$ similarly $\Lambda - \alpha_j \in S_\rho$. We have further $f_i(\Lambda - \alpha_j) > 0$ and $f_j(\Lambda - \alpha_i) > 0$ and for $k \neq j$, we have $f_k(\Lambda - \alpha_j) > 0$ (Lemma 6). It follows that

$$f_k(\mu_\rho) \geqq f_k(\Lambda - \alpha_i + \Lambda - \alpha_j) > 0$$

for every k with $1 \leqq k \leqq i$. Hence in this case also $f_i(\mu_\rho) > 0$ for $1 \leqq i \leqq l$.

If case (i) and case (ii) do not hold we have for every i with $1 \leqq i \leqq 1$ and $\langle \Lambda, \alpha_i \rangle > 0$, $f_i(\Lambda) \leqq 1$ with equality occurring for almost one i. Let $\mathfrak{g} = \sum_s \mathfrak{g}_s$ be the decomposition of $\mathfrak{g}$ into its simple components. Then $\Lambda = \sum_s \Lambda_s$ where Λ_s are the weights of representations of $\mathfrak{g}_s$ (we mean by this a representation which is trivial on every component except $\mathfrak{g}_s$). From $f_k(\Lambda) \geqq f_k(\Lambda_s)$ for every s, it follows that each $\mathfrak{g}_s$ must be one of the classical Lie algebras of Lemma 7 and Λ_s the highest weight of the corresponding natural representation. Further we cannot have more than one s with $\mathfrak{g}_s$ of the types of B_n, C_n $(n \geqq 2)$, D_n; since if this happened, i. e., say $\mathfrak{g}_{s_1}$, and $\mathfrak{g}_{s_2}$ are both of one of these types, then we have for some k_1, k_2, $f_k(\Lambda_{s_1}) = f_{k_2}(\Lambda_{s_2}) = 1$, (see case ii). Hence we have at most one component $\mathfrak{g}_r$ of the type $B_n C_n$ or

D_n. It follows that $\mathfrak{g}=\sum_s \mathfrak{g}_s$ where $\mathfrak{g}_s$ for all but at most one s say s_0 is an algebra of type A_n and $\mathfrak{g}_{s_0}$ can be any one of A_n, B_n, C_n or D_n and ρ is the tensor product of the corresponding natural representation or their duals.

Definition 1. A representation ρ of a semisimple algebra $\mathfrak{g}$ is of *type P* if $f_i(\mu_\rho) > 0$ for $1 \leqq i \leqq l$.

LEMMA 11. *Let $\mathfrak{g}$ be the three dimensional Lie algebra and ρ a finite dimensional representation in a vector space V. Let $v \in V$ be a weight vector of ρ (w.r.t. $\mathfrak{h}$) the corresponding weight being μ. Let $\pm\alpha$ be the roots of $\mathfrak{g}$ and E'_α, $E'_{-\alpha}$ be root vectors of $\mathfrak{g}$ such that $A(E'_\alpha, E'_{-\alpha})=1$ where A is an invariant bilinear form on $\mathfrak{g}$ which is positive definite on $\mathfrak{h}^*$. Let further $\langle\ ,\ \rangle$ be a hermitian positive definite scalar product on V with respect to which V admits an orthonormal decomposition into irreducible components. Then we have*

$$\langle\rho(E'_\alpha E'_{-\alpha}+E'_{-\alpha}E'_\alpha)(v),v\rangle \geqq |A(\mu,\alpha)|\langle v,v\rangle.$$

(A induces a positive definite scalar product on the dual of $\mathfrak{h}^*$.) If moreover $E'_\alpha(v)\neq 0$, we have $\langle\rho(E'_\alpha E'_{-\alpha}+E'_{-\alpha}E'_\alpha)(v),v\rangle > 0$.

Proof. We may clearly assume that ρ is irreducible.

Let Λ be the highest weight of ρ. If $H \in \mathfrak{h}^*$ is such that $A(H,H)=1$, then we have from Lemma 4, that

$$\rho(E'_\alpha E'_{-\alpha}+E'_{-\alpha}E'_\alpha+H^2)(v)=A(\Lambda,\Lambda)+A(\Lambda,\alpha)$$

(α the positive root). It follows that if $\langle v,v\rangle=1$

$$\begin{aligned}\langle\rho(E'_\alpha E'_{-\alpha}+E'_\alpha E'_{-\alpha})(v),v\rangle &= A(\Lambda,\Lambda)+A(\Lambda,\alpha)\,\mu(H)^2\\ &= A(\Lambda,\Lambda)+A(\Lambda,\alpha)-A(\mu,\mu).\end{aligned}$$

so that

$$\langle\rho(E'_\alpha E'_{-\alpha}+E'_{-\alpha}E'_\alpha)(v),v\rangle \geqq A(\Lambda,\alpha),$$

since $A(\Lambda,\Lambda)\geqq A(\mu,\mu)$, Λ being the highest weight. Once again since Λ is the highest weight $A(\Lambda,\alpha)\geqq|A(\mu,\alpha)|$. Hence the assertion; also if $E'_\alpha v\neq 0$, ρ is non-trivial so that $A(\Lambda,\alpha)>0$ which proves the last assertion.

LEMMA 12. *Let $\mathfrak{g}$ be a reductive Lie algebra, $\mathfrak{c}$ its centre and $\mathfrak{g}'$ its semisimple part. Let ρ be an irreducible finite dimensional representation of $\mathfrak{g}$. Let $\mathfrak{h}'$ be a Cartan subalgebra of $\mathfrak{g}'$, and $\mathfrak{h}=\mathfrak{h}'\oplus\mathfrak{c}$ that of $\mathfrak{g}$. Then we have*

a) *If v is a weight vector for $\rho|_{\mathfrak{g}'}$, then it is a common eigenvector for the whole of $\mathfrak{h}$*

b) *$\rho\,|_{\mathfrak{g}'}$ is irreducible*

c) *If α is a simple root of $\mathfrak{g}'$ we denote by α also the extension to $\mathfrak{h}$ defined by $\alpha(\mathfrak{c}) = 0$; let v_0 be the highest weight of $\rho\,|_{\mathfrak{g}'}$ and Λ_0 the corresponding linear form on $\mathfrak{h}$* (see (a)); *then for any weight vector v of $\rho\,|_{\mathfrak{g}'}$ we have $\rho(H)(v) = \Lambda(H)(v)$ for $H \in \mathfrak{h}$, Λ a linear form, $\Lambda = \Lambda_0 - \sum m_i\alpha_i$, $m_i \geqq 0$ integers and α_i simple roots of $\mathfrak{g}$.*

Proof. a) is a consequence of Schur's lemma: the whole of V is an eigen space of $\rho(H)$ for every $H \in \mathfrak{c}$.

b) is a consequence of Schur's lemma and the Burnside theorem: the associative algebra generated by $\rho(\mathfrak{g}')$ in End V is the same as that generated by $\rho(\mathfrak{g})$ viz. the whole of End (V).

c) is a consequence of a) and b) and the corresponding fact for irreducible representations of a semisimple Lie algebra.

LEMMA 13. *Let $\mathfrak{g}$ be a semisimple Lie algebra and Λ a weight of a finite dimensional representation ρ. Suppose for a root $\alpha \in \Delta$, $\langle\Lambda, \alpha\rangle_{\mathfrak{g}} > 0$, then $\Lambda - \alpha$ is a weight of ρ.*

Proof. In fact the entire "α-series" $\Lambda, \Lambda - \alpha, \cdots, \Lambda - \dfrac{2\langle\Lambda, \alpha\rangle}{\langle\alpha, \alpha\rangle}\cdot\alpha$ are weight of ρ. See, for instance, [7], exposé 19, Theorem 1.

Definition 1′. A representation ρ of a reductive Lie algebra $\mathfrak{g}$ is of type P if $\{\rho\,|\,(\text{semisimple part of } \mathfrak{g})\}$ is of type P (See Definition 1.). (If $[\mathfrak{g}, \mathfrak{g}] = 0$, no representation is of type P).

3. Complexification of a real simple Lie algebra. In this section $\mathfrak{g}_0$ will be a *real simple noncompact Lie algebra* and $\mathfrak{g}$ will denote the complexification of $\mathfrak{g}_0$. We identify complexifications of real subspaces of $\mathfrak{g}_0$ with the corresponding subspaces of $\mathfrak{g}$. Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be a Cartan-decomposition of $\mathfrak{g}_0$, $\mathfrak{k}_0$ being the algebra. Let $\mathfrak{k}$ (resp. $\mathfrak{p}$) be the complexification of $\mathfrak{k}_0$ (resp. $\mathfrak{p}_0$). Let θ_0 be the involution of $\mathfrak{g}_0$ correspondind to the above Cartan-decomposition and θ the extension of θ_0 to $\mathfrak{g}$. Let $\mathfrak{h}_{\mathfrak{k}_0}$ be a Cartan subalgebra of $\mathfrak{k}_0$. Let $\mathfrak{h}_0$ be a Cartan subalgebra of $\mathfrak{g}_0$ such that $\mathfrak{h}_0 \supset \mathfrak{h}_{\mathfrak{k}_0}$. Let $\mathfrak{h}$ (resp. $\mathfrak{h}_{\mathfrak{k}}$) denote the complexification of $\mathfrak{h}_0$ (resp. $\mathfrak{h}_{\mathfrak{k}_0}$).

LEMMA 14. $\mathfrak{h} = \mathfrak{h}_{\mathfrak{k}} \oplus \mathfrak{h} \cap \mathfrak{p}$. *Consequently $\mathfrak{h}$ is stable under θ.*

Proof. Let $H \in \mathfrak{h}$; then $H = H_1 + H_2$ where $H_1 \in \mathfrak{k}$, $H_2 \in \mathfrak{p}$. Now for $H' \in \mathfrak{h}_{\mathfrak{k}_0}$, we have $0 = [H, H'] = [H_1, H'] + [H_2, H']$; now $[H_1, H'] \in \mathfrak{k}$ and $[H_2, H'] \in \mathfrak{p}$ so that each of them is zero; hence $H_1 \in \mathfrak{h}_{\mathfrak{k}}$. Hence the lemma.

Let Δ be the root system of $\mathfrak{g}$ w.r.t. $\mathfrak{h}$. We choose for every $\alpha \in \Delta$ a root vector E_α of the root α such that $\langle E_\alpha, E_{-\alpha}\rangle_{\mathfrak{g}} = 1$ let $H_\alpha = [E_\alpha, E_{-\alpha}]$; then $H_\alpha \in \mathfrak{h}$ and let $\mathfrak{h}^* = \sum_{\alpha \in \Delta} \boldsymbol{R} H_\alpha$. Also for $\alpha \in \Delta$, the linear form α^θ defined by $\alpha^\theta(H) = \alpha(\theta(H))$ for $H \in \mathfrak{h}$ (see Lemma 14), is again in Δ: in fact for $H \in \mathfrak{h}$, we have

$$[H, \theta(E_\alpha)] = \theta[\theta(H), E_\alpha] = \alpha(\theta(H))\theta(E_\alpha) = \alpha^\theta(H) \cdot \theta(E_\alpha).$$

We assume E_α to be so chosen that $\theta(E_\alpha) = E_{\alpha^\theta}$: this is not in violation of the earlier condition since $\langle \theta(X), \theta(Y)\rangle_{\mathfrak{g}} = \langle X, Y\rangle_{\mathfrak{g}}$, θ being an automorphism of the Lie algebra structure. With this notation, we have the following

LEMMA 15. a) *If $\alpha \in \Delta$, $\alpha = \alpha^\theta$ if and only if $\alpha(H) = 0$ for $H \in \mathfrak{h}_{\mathfrak{p}}$; also if $\alpha = \alpha^\theta$, then $\theta(E_\alpha) = \pm E_\alpha$.* ($\mathfrak{h}_{\mathfrak{p}} = \mathfrak{h} \cap \mathfrak{p}$)

b) $\mathfrak{h}^* = i\mathfrak{h}_{\mathfrak{k}_0} \oplus \mathfrak{h}_{\mathfrak{p}_0}$ ($\mathfrak{h}_{\mathfrak{p}_0} = \mathfrak{h} \cap \mathfrak{p}_0$); *as a real vector space the rank of $\mathfrak{h}^*$ is the same as that of $\mathfrak{h}$ over $\boldsymbol{C}$.*

Proof. (a) follows from Lemma 14 and the fact that $\theta(H) = -H$ for $H \in \mathfrak{h}_{\mathfrak{p}}$; if $\alpha = \alpha^\theta$, then $\theta(E_\alpha)$ being again a root vector of root $\alpha^\theta = \alpha$, we have $\theta(E_\alpha) = \lambda E_\alpha$, $\lambda \in \boldsymbol{C}$; since $\theta^2 = 1$, $\lambda = \pm 1$.

(b) is a consequence of the fact that $\mathfrak{k}_0 + i\mathfrak{p}_0$ is a compact real form of $\mathfrak{g}$. (See [3], p. 219, Lemma 3.1.). Hence the lemma.

Every real valued linear form on $\mathfrak{h}^*$ extends uniquely as a complex linear form on $\mathfrak{h}$; we identify each such form with its unique extension: this defines an identification of the real dual of $\mathfrak{h}^*$ and the real subspace of $\boldsymbol{C}$-linear forms on $\mathfrak{h}$ which are real on $\mathfrak{h}^*$. On the dual of the space $\mathfrak{h}^*$ we introduce a lexicographic order as follows. The algebra $\mathfrak{k}_0$ is reductive; hence $\mathfrak{k}_0 = \mathfrak{c}_0 + \mathfrak{k}'_0$ where $\mathfrak{k}'_0$ is the semisimple part of $\mathfrak{k}_0$ and $\mathfrak{c}_0$ is the centre of $\mathfrak{k}_0$. The centre $\mathfrak{c}$ can be zero; if it is not, it is one dimensional. Consequently we have $\mathfrak{h}_{\mathfrak{k}_0} = \mathfrak{h}_{\mathfrak{k}'_0} \oplus \mathfrak{c}_0$ so that $\mathfrak{h}^* = i\mathfrak{c}_0 \oplus i\mathfrak{h}_{\mathfrak{k}'_0} \oplus \mathfrak{h}_{\mathfrak{p}_0}$. Let H_0, $H_1, \cdots, H_l$ be an orthonormal basis of $\mathfrak{h}^*$ with respect to the Killing form so chosen that $H_0 \in i\mathfrak{c}_0$, $\{H_i\}_{1 \le i \le p}$ is an (orthonormal) basis of $i\mathfrak{h}_{\mathfrak{k}'_0}$ and for $i > p$, $H_i \in \mathfrak{h}_{\mathfrak{p}_0}$ (necessarily). (When $\mathfrak{c}_0 = 0$, we set $H_0 = 0$). Let α, β be (real) linear forms on $\mathfrak{h}^*$.

Definition 2. $\alpha > \beta$ if the first non-vanishing difference $(\alpha - \beta)(H_i)$ (as i runs from 0 to l) is greater than zero.

Let $\Delta^+ = \{\alpha \mid \alpha \in \Delta, \alpha > 0\}$. Let

$$A = \{\alpha \mid \alpha \in \Delta, \theta(E_\alpha) = E_\alpha\},\ B = \{\alpha \mid \alpha \in \Delta, \alpha \neq \alpha^\theta\}$$

$$\text{and } C = \{\alpha \mid \alpha \in \Delta, \theta(E_\alpha) = -E_\alpha\},$$

and $A^+ = A \cap \Delta^+$, $B^+ = B \cap \Delta^+$, $C^+ = C \cap \Delta^+$. Also, for a linear form λ on $\mathfrak{h}$ we denote its restriction to $\mathfrak{h}_{\mathfrak{k}}$ by $r(\lambda)$. Finally, let σ denote the adjoint representation of $\mathfrak{k}$ in $\mathfrak{p}$. We have then

LEMMA 16.

a) $\mathfrak{k} = \mathfrak{h}_{\mathfrak{k}} \oplus \{ \sum_{\alpha \in A^+} (\boldsymbol{C}E_\alpha \oplus \boldsymbol{C}E_{-\alpha})\} \oplus \sum_{\alpha \in B^+} \boldsymbol{C}(E_\alpha + \theta(E_\alpha)) \oplus \boldsymbol{C}(E_{-\alpha} + \theta(E_{-\alpha}))\}$

b) $\mathfrak{p} = \mathfrak{h}_{\mathfrak{p}} \oplus \{ \sum_{\alpha \in C^+} (\boldsymbol{C}E_\alpha \oplus E_{-\alpha})\} \oplus \sum_{\alpha \in B^+} \boldsymbol{C}(E_\alpha - \theta(E_\alpha)) \oplus \boldsymbol{C}(E_{-\alpha} - \theta(E_{-\alpha}))\}$.

Consequently,

c) *If* $\alpha \in A^+ \cup B^+$, $r(\alpha)$ *is a root of* $\mathfrak{k}$ *with* $E_\alpha + \theta(E_\alpha)$ *as a root vector.*

d) *if* $\alpha \in B^+ \cup C^+$, *then* $r(\alpha)$ *is a weight of* σ *with* $E_{\alpha^\theta} - \theta(E_{-\alpha})$ *as a weight vector.*

e) *If* $\alpha, \beta \in \Delta$ *are such that* $r(\alpha) = r(\beta)$ *then* $\alpha = \beta$ *or* $\alpha = \beta^\theta$; *consequently in* (c) (resp. (d)) *the root vector* (resp. *the weight vector*) *is unique up to a scalar multiple.*

f) *For* $\alpha \in \Delta$, $r(\alpha) \neq 0$.

Proof. a) and b) are immediate from the fact that $x \to \dfrac{x + \theta(x)}{2}$ and $x \to \dfrac{x - \theta(x)}{2}$ are projections of $\mathfrak{g}$ onto $\mathfrak{k}$ and $\mathfrak{p}$ respectively

Proof of c). In fact we have for $H \in \mathfrak{h}_{\mathfrak{k}}$,

$$[H, E_\alpha + \theta(E_\alpha)] = \alpha(H) \cdot E_\alpha + \alpha^\theta(H)\,\theta(E_\alpha)$$

and since for $H \in \mathfrak{h}_{\mathfrak{k}}$, $\alpha^\theta(H) = \alpha(\theta(H)) = \alpha(H)$, we have

$$[H, E_\alpha + \theta(E_\alpha)] = \alpha(H)(E_\alpha + \theta(E_\alpha)) = r(\alpha)(H)(E_\alpha + \theta(E_\alpha)).$$

Hence (c).

Proof of d) is analogous to that of (c).

Proof of e). If $\alpha, \beta \in A^+ \cup C^+$, then since $\alpha(\mathfrak{h}_{\mathfrak{p}}) = \beta(\mathfrak{h}_{\mathfrak{p}}) = 0$, $r(\alpha) = r(\beta)$ implies $\alpha = \beta$. Suppose next that $\alpha, \beta \in A^+ \cup B^+$; then $r(\alpha)$, $r(\beta)$ being the same root of $\mathfrak{k}$, $E_\alpha + \theta(E_\alpha)$ and $E_\beta + \theta(E_\beta)$ are linearly dependent; hence either E_α and E_β coincide or E_α and $\theta(E_\beta) = E_{\beta^\theta}$ coincide. Hence $\alpha = \beta$ or $\alpha = \beta^\theta$. Suppose next that $\alpha \in B^+$, $\beta \in C^+$; then since $\beta(\mathfrak{h}_{\mathfrak{p}}) = 0$ and $r(\alpha) = r(\beta)$ we have $\langle \alpha, \beta \rangle_{\mathfrak{g}} = \langle \beta, \beta \rangle_{\mathfrak{g}} > 0$. It follows that $\alpha - \beta$ is a root of $\mathfrak{g}$ (see Lemma 13); now $(\alpha - \beta)^\theta \neq (\alpha - \beta)$ and so $r(\alpha - \beta)$ is root of $\mathfrak{k}$; on the other hand clearly $r(\alpha - \beta) = 0$, a contradiction.

Proof of f). This follows from the fact that for $\alpha \in A^+ \cup C^+$, $\alpha(\mathfrak{h}_\mathfrak{p}) = 0$ so that if $\alpha(\mathfrak{h}_\mathfrak{k}) = 0$, $\alpha(\mathfrak{h}) = 0$; if $\alpha \in B^+$, $r(\alpha)$ is a root of $\mathfrak{k}$ hence $r(\alpha) \neq 0$.

This completes the proof of Lemma 16. We have also the following

LEMMA 17. *If* $\alpha \neq \alpha^\theta$, $\frac{1}{2}\langle E_\alpha + \theta(E_\alpha), E_{-\alpha} + \theta(E_{-\alpha})\rangle = 1$ *and*

$$[E_\alpha + \theta(E_\alpha), E_{-\alpha} + \theta(E_{-\alpha})] = H_\alpha + H_{\alpha^\theta}.$$

Proof. If $\alpha = -\alpha^\theta$ then $\alpha(\mathfrak{h}_k) = 0$ contradicting (f) of Lemma 16. Hence $\alpha^\theta \neq -\alpha$; it follows that $\langle E_\alpha, E_{\alpha^\theta}\rangle = 0$ for $\alpha \neq \alpha^\theta$, i.e., $\langle E_\alpha, \theta(E_\alpha)\rangle = 0$. Hence the first assertion. To prove the second, it is sufficient to show that $[E_\alpha, \theta(E_{-\alpha})] = 0$ if $\alpha \neq \alpha^\theta$; in fact, if $[E_\alpha, \theta(E_{-\alpha})]$ is non-zero, $\alpha - \alpha^\theta$ is a root of $\mathfrak{g}$; but then $r(\alpha - \alpha^\theta) = 0$ contradicting (f) of Lemma 16. Hence the lemma.

LEMMA 18. *If* $\mathfrak{k}$ *is semisimple, then for* $\alpha \in \Delta$, $r(\alpha)$ *is a half integer linear combination of simple hoots of* $\mathfrak{k}$; *if moreover* $B \neq \emptyset$, *then* $r(\alpha)$ *is actually an integral linear combination of simple roots of* $\mathfrak{k}$.

Proof. Let $\gamma_1, \cdots, \gamma_p$ be the simple roots of $\mathfrak{k}$. If $\alpha \in A^+ \cup B^+$ then $r(\alpha)$ is a root of $\mathfrak{k}$ and there is nothing to prove. Let $\beta \in B^+ \cup C^+$ be the root such that $r(\beta)$ is the highest weight of the adjoint representation σ of $\mathfrak{k}$ in $\mathfrak{p}$. Then for $\alpha \in B^+ \cup C^+$, we have $r(\alpha) = r(\beta) - \sum_{i=1}^{p} m_i\gamma_i$ where m_i are non-negative integers. It is therefore sufficient to prove that $r(\beta) = \sum n_i\gamma_i$ where n_i are half integers. Nof if $\beta \in B \cup C$ so is $-\beta$. Hence

$$r(-\beta) = r(\beta) - \sum_{i=1}^{p} m_i\gamma_i$$

where m_i are integers. Hence the first assertion. From the proof it is clear that if for some one $\alpha \in B^+ \cup C^+$, $r(\alpha)$ is an integral linear combination of the γ_i, then the same is true of every $\alpha \in B^+ \cup C^+$; if now $B^+ \neq \emptyset$, then for $\alpha \in B^+$, $r(\alpha)$ is a root and this proves the second assertion.

The following result is well known.

LEMMA 19. *We use the lexicographic order on the dual of* $\mathfrak{h}^*$ *defined earlier. If* $\mathfrak{k}$ *is not semisimple then* $\mathfrak{k} = \mathfrak{k}' + \mathfrak{c}$ *where* $\mathfrak{k}'$ *is semisimple and* $\mathfrak{c}$ *is the* 1-*dimensional centre. Also* $\mathfrak{h}_\mathfrak{k} = \mathfrak{h}$, *and* $B^+ = \emptyset$. *Every simple root* α *of* $\mathfrak{k}$ *is also a simple root of* $\mathfrak{g}$, *so that the system of simple roots of* $\mathfrak{g}$ *is* $\alpha_1, \cdots, \alpha_{l-1}, \alpha_l$ *where* $\alpha_1, \cdots, \alpha_{l-1}$ *are simple roots of* $\mathfrak{k}$ *and* $\alpha_l \in C^+$. *Moreover for* $\alpha \in C^+$, $\alpha(H_0) = \pm \alpha_l(H_0)$. *Finally,* $-\alpha_l|_{\mathfrak{h}_{\mathfrak{k}'}}$ *is a highest weight of a component of* $\sigma|_{\mathfrak{k}'}$. *The highest root* β $(\neq \alpha_l$ *if rank* $\mathfrak{g} > 1)$ *is another of the highest weights of* σ.

Proof. That $\mathfrak{k} = \mathfrak{k}' + \mathfrak{c}$ where $\mathfrak{c}$ is one dimensional is well known as also the fact that rank $\mathfrak{g} =$ rank $\mathfrak{k}$ i. e., $\mathfrak{h}_{\mathfrak{k}} = \mathfrak{h}$; hence $B^+ = \emptyset$.

A root $\alpha \in \Delta^+$ is a root of $\mathfrak{k}$ if and only if $\alpha(H_0) = 0$; this follows from the fact that an element $X \in \mathfrak{g}$ commutes with H_0 if and only if $X \in \mathfrak{k}$. If then α is a simple root of $\mathfrak{k}$ and $\alpha = \beta_1 + \gamma$ where $\beta_1, \gamma \in \Delta^+$, then since $\beta_1(H_0) \geqq 0$, $\gamma(H_0) \geqq 0$, $\beta_1(H_0) = \gamma(H_0) = 0$ i. e., β_1, γ are roots of $\mathfrak{k}$, a contradiction. Now $\mathfrak{k}'$ being of rank $l - 1$, let $\alpha_1, \cdots, \alpha_{l-1}$ be the simple roots of $\mathfrak{k}$. Let α_l be the one other simple root of $\mathfrak{g}$. Since $(\alpha_i)_{1 \leq i \leq l}$ are linearly independent and $\alpha_i(H_0) = 0$ if $i < l$, $\alpha_l(H_0) > 0$ $(\alpha_l > 0)$.

The last two assertions are consequences of the fact that $\sigma(E_\alpha)(E_\beta) = 0$ for every $\alpha \in A^+$ and $[E_{\alpha_i}, E_{-\alpha_l}] = 0$ if $i < l$ (see [7], exposé 10).

LEMMA 20. *Let ρ be an irreducible non-trivial representation of $\mathfrak{g}$. Suppose now that $\mathfrak{k}$ is not semisimple. Then if Λ is the highest weight of ρ, $\Lambda(H_0) > 0$.*

Proof. Since $\mathfrak{k}$ is not semisimple, $\mathfrak{g}$ is simple (see [3]). It follows that ρ is faithful. Consequently we have $\Lambda = \sum_{i=1}^{l} m_i\alpha_i$ with $m_i > 0$ (in the notation of the previous lemma) (see Lemma 6). We have from Lemma 19 above, that $\alpha_i(H_0) = 0$ for $i < l$ so that $\Lambda(H_0) = m_l\alpha_l(H_0) > 0$. This proves the lemma.

LEMMA 21. *When $\mathfrak{k}$ is semisimple, σ is faithful and irreducible; if $\mathfrak{k}$ is not, $\sigma = \sigma_1 \oplus \sigma_2$ where both σ_1 and σ_2 are faithful and irreducible.*

Proof. That σ is faithful is well known as also that it is irreducible when $\mathfrak{k}$ is semisimple and that if $\mathfrak{k}$ is not semisimple $\sigma = \sigma_1 \oplus \sigma_2$ *i. e.*, σ has only two irreducible components σ_1 and σ_2. Now $\sigma_1 |_{\mathfrak{k}_0}$ and $\sigma_2 |_{\mathfrak{k}_0}$ are equivalent to σ_0 as real representations. Hence these are faithful representations of $\mathfrak{k}_0$. Suppose now that $\mathfrak{k}_0 = \mathfrak{c}_0 \oplus \sum_i \mathfrak{k}_0{}^i$ where $\mathfrak{k}_0{}^i$ are simple, then $\mathfrak{k} = \mathfrak{c} + \sum_i \mathfrak{k}^i$, where $\mathfrak{k}_i$ are complex simple algebras. If σ_1 (or σ_2) is zero on $\mathfrak{k}^i$ then it is zero on $\mathfrak{k}_0{}^i$, a contradiction. As for $X \in \mathfrak{c}_0$, $\sigma_i(X)$ is a non-trivial scalar operator if $X \neq 0$. This proves the lemma.

LEMMA 22. *If $\alpha \in A^+$ is a simple root of $\mathfrak{g}$, then $r(\alpha)$ is a simple root of $\mathfrak{k}$.*

Proof. In the case when rank $\mathfrak{g} =$ rank $\mathfrak{k}$, every positive root of $\mathfrak{k}$ is a positive root of $\mathfrak{g}$ and the lemma is immediate. Suppose then that rank $\mathfrak{g}$ $>$ rank $\mathfrak{k}$ and, if possible, let $r(\alpha) = r(\beta) + r(\gamma)$ where α is a simple root of $\mathfrak{g}$ and $r(\beta) > 0$, $r(\gamma) > 0$ are roots of $\mathfrak{k}$ i. e. $\beta, \gamma \in B^+ \cup A^+$. It follows

then that $[E_\beta+\theta(E_\beta),E_\gamma+\theta(E_\gamma)]\neq 0$ (hence a scalar multiple of E_α). In particular we have either $[E_\beta,E_\gamma]\neq 0$ or $[\theta(E_\beta),E_\gamma]\neq 0$ so that either $\beta+\gamma$ or $\beta^\theta+\gamma$ is a root. Let $\beta+\gamma$ be a root of $\mathfrak{g}$ (we may clearly assume this without loss of generality). Since $r(\beta+\gamma)=r(\alpha)$, it follows from (e) of Lemma 16, that $\alpha=\beta+\gamma$ or $\alpha^\theta=(\beta+\gamma)$: in either case α is not a simple root of $\mathfrak{g}$. This proves the lemma.

LEMMA 23. *Let $\mathfrak{g}_0$ be a real non-compact simple Lie algebra and ρ a non-trivial irreducible representation of $\mathfrak{g}=\mathfrak{g}_0\otimes \boldsymbol{C}$ such that*

i) *the adjoint representation σ of $\mathfrak{k}$ in $\mathfrak{p}$ is not of type P* (see Definition 1′ (below Lemma 13, § 2)).

ii) *if Λ is the highest weight of ρ, then for every $\alpha\in C^+\cup B^+$ such that $r(\alpha)$ is not a highest weight of σ, $\langle\Lambda,\alpha\rangle_{\mathfrak{g}}=0$. Then $\mathfrak{g}_0\cong\mathfrak{so}_0(n,1)$ or $\mathfrak{su}(n,1)$ and $\Lambda=m\cdot\mu_N$ where m is a positive integer and μ_N is the highest weight of the corresponding natural representation of $\mathfrak{g}\cong\mathfrak{so}(n+1,\boldsymbol{C})$ or $\mathfrak{sl}(n+1,\boldsymbol{C})$ as the case may be.*

Proof. We divide the proof into three parts. (**A**) $\operatorname{Rank}\mathfrak{g}>\operatorname{Rank}\mathfrak{k}$, (**B**) $\operatorname{Rank}\mathfrak{g}=\operatorname{Rank}\mathfrak{k}$; $\mathfrak{k}$ semisimple, (**C**) $\mathfrak{k}$ non-semisimple (hence $\operatorname{rank}\mathfrak{g}=\operatorname{rank}\mathfrak{k}$).

Case **A**. *Rank* $\mathfrak{g}>$ *Rank* $\mathfrak{k}$. In view of Lemma 10, we see that $\mathfrak{k}$ $\mathfrak{k}_1\times\mathfrak{k}_2\times\mathfrak{k}$ where $\mathfrak{k}_1$ and $\mathfrak{k}_2$ are products of algebras of the type $\mathfrak{sl}(n,\boldsymbol{C})$ and $\mathfrak{k}_3$ is one of the algebras $\mathfrak{sl}(n,\boldsymbol{C})$, $\mathfrak{so}(n,\boldsymbol{C})$ $(n\neq 4)$ or $\mathfrak{sp}(n,\boldsymbol{C})$; moreover $\rho=\rho_1\otimes\rho_2\otimes\rho_3$ where ρ_1 (resp. ρ_2, resp. ρ_3) is the representation of $\mathfrak{g}_1$ (resp. $\mathfrak{g}_2$, resp. $\mathfrak{g}_3$) obtained by forming the tensor product of the natural (resp. dual of the natural, resp. natural) representation of its simple components. For the algebras $\mathfrak{sl}(n,\boldsymbol{C})$, $\mathfrak{sp}(n,\boldsymbol{C})$, $\mathfrak{so}(2n,\boldsymbol{C})$, the expression for the highest weight of the natural representation in terms of the simple roots of $\mathfrak{k}$ will involve non-integral coefficients. Hence (Lemma 18) $\mathfrak{k}$ is an algebra of type B_n and ρ is the natural representation, i.e., $\mathfrak{k}=\mathfrak{so}(2n+1,\boldsymbol{C})$ and ρ is the natural representation. It follows that $\mathfrak{k}_0=\mathfrak{so}(2n+1)$ and that σ_0 is the natural representation (this follows from the fact that if V, V' are real vector spaces which are simple $\mathfrak{k}_0$-modules, then $V\cong V'$ (as $\mathfrak{k}_0$-modules) if and only if $V\otimes_{\boldsymbol{R}}\boldsymbol{C}\cong V'\otimes_{\boldsymbol{R}}\boldsymbol{C}$). Now the algebra $\mathfrak{so}_0(2n+1,1)$ has $\mathfrak{so}(2n+1)$ for a maximal compactly imbedded algebra and the representation of this maximal compact algebra in its supplement in $\mathfrak{so}_0(2n+1,1)$ (w.r.t. the Killing form) is clearly the natural representation. Hence by the uniqueness theorem (Theorem 3.6, page 337, [3]) $\mathfrak{g}_0=\mathfrak{so}(2n+1,1)$. This proves part of the Lemma in Case **A**.

Case **B.** *Rank* $\mathfrak{g}$ = Rank $\mathfrak{k}$; $\mathfrak{k}$ *semisimple.* Once again combining Lemma 18 and Lemma 10 and also looking at the coefficients of the simple roots of $\mathfrak{k}$ in the expression for the highest weights (Lemma 7), we see that only one of the algebras $\mathfrak{sl}(2,\boldsymbol{C})$, B_n, C_n, D_n are admissible as components of $\mathfrak{k}$. (Lemma 18 asserts that the coefficients of the simple roots of $\mathfrak{k}$ in the expression for the highest weight $r(\beta)$ are half-integers).

Let then $\mathfrak{k} = \mathfrak{k}_1 \times \mathfrak{k}_2 \cdots \times \mathfrak{k}_m$. Let β be the highest weight of σ and γ_i the *unique* simple root of $\mathfrak{k}_i$ (considered as a root of $\mathfrak{k}$) such that $\langle\beta,\gamma_i\rangle_{\mathfrak{k}} \neq 0$ (Lemma 10). Let $\delta_1,\cdots,\delta_n$ be the rest of the simple roots of $\mathfrak{k}$. (We have written β for $r(\beta)$ since $\mathfrak{h}_{\mathfrak{k}} = \mathfrak{h}$; we will continue to do so in the rest of the proof). We have then

$$\beta = \sum_{i=1}^{m} q_i\gamma_i + \sum_{j=1}^{n} r_j\delta_j \tag{I}$$

where $\frac{1}{2} \leqq q_i \leqq 1$ and $r_j > 0$ (Lemma 7 and the form of the highest weights). We have here moreover that the q_i and the r_j are half integers. Now, $\langle\ ,\ \rangle_{\mathfrak{g}}$ restricted to $\mathfrak{k}$ is a scalar product on $\mathfrak{k}$ which is clearly invariant under $\mathfrak{k}$. Also, since for $\alpha \in A^+$, $E_\alpha \in \mathfrak{k}$, it follows that, if H'_α is the unique element of $\mathfrak{h}_k$ with $\langle H'_\alpha, H\rangle_{\mathfrak{k}} = \alpha(H)$, then $H'_\alpha \in \mathfrak{h}^*$ so that $\mathfrak{h}^* = \sum_{\alpha\in A^+} \boldsymbol{R}H'_\alpha$. Hence $\langle\ ,\ \rangle_{\mathfrak{g}}$ is positive definite on $\sum_{\alpha\in A^+} \boldsymbol{R}H'_\alpha$. Hence by Lemma 1, since $\langle\beta,\gamma_i\rangle_{\mathfrak{k}} > 0$, $\langle\beta,\gamma_i\rangle_{\mathfrak{g}} > 0$. It follows then that $\beta - \gamma_i$ is a root of $\mathfrak{g}$ (Lemma 13). Now γ_i is a root of $\mathfrak{k}$ as also a root of $\mathfrak{g}$. Hence $\gamma_i \in A^+$. It follows that $\beta - \gamma_i \in C$ so that $\langle\Lambda,\beta-\gamma_i\rangle_{\mathfrak{g}} = 0$ (condition ii), i.e., $\langle\Lambda,\beta\rangle_{\mathfrak{g}} = \langle\Lambda,\gamma_i\rangle_{\mathfrak{g}}$. Now from the expression for β ((I), above), we see then that

$$\langle\Lambda,\beta\rangle_{\mathfrak{g}} = \sum_{i=1}^{m} q_i\langle\Lambda,\gamma_i\rangle_{\mathfrak{g}} + \sum_{j=1}^{n} r_j\langle\Lambda,\delta_j\rangle_{\mathfrak{g}} \cdots \tag{II}$$

Now since Λ is the highest weight, $\langle\Lambda,\delta_j\rangle_{\mathfrak{g}} \geqq 0$: δ_j is a positive root of $\mathfrak{g}$. Substituting $\langle\Lambda,\beta\rangle_{\mathfrak{g}} = \langle\Lambda,\gamma_i\rangle_{\mathfrak{g}}$ in the above, we obtain,

$$1 \geqq \sum_{i=1}^{m} q_i.$$

Since q_i are half integers, this leaves us three possibilities:

(a) $m = 1; q_1 = \frac{1}{2}$, (b) $m = 1, q_1 = 1$, (c) $m = 2, q_1 = q_2 = \frac{1}{2}$.

In case (c), clearly $\mathfrak{k} = \mathfrak{sl}(2,\boldsymbol{C}) \times \mathfrak{sl}(2,\boldsymbol{C})$ or what is the same $\mathfrak{so}(4)$ and σ is the natural representation. It follows as before, that $\mathfrak{g}_0 = \mathfrak{so}(4,1)$.

In case (a), $\mathfrak{k} = \mathfrak{sl}(2,\boldsymbol{C})$ i.e. $\mathfrak{k}$ is of rank 1; but then rank $\mathfrak{g} = 1$; $\mathfrak{g} = \mathfrak{k} = \mathfrak{sl}(2,\boldsymbol{C})$ contradicting the non-compactness of $\mathfrak{g}_0$.

In case (b), a priori, $\mathfrak{k}$ could be either $\mathfrak{sp}(n, \boldsymbol{C})$ or $\mathfrak{so}(2n+1, \boldsymbol{C})$ or $\mathfrak{so}(2n, \boldsymbol{C})$. In the former two cases there is a root α of $\mathfrak{k}$ such that a rational multiple of α (α in the case of $\mathfrak{so}(2n+1, \boldsymbol{C})$ and $\alpha/2$ for $\mathfrak{sp}(n, \boldsymbol{C})$) are weights of σ (i.e. of the natural representation). On the other hand since rank $\mathfrak{k} =$ Rank $\mathfrak{g}$, every weight of σ and every root of $\mathfrak{k}$ are roots of $\mathfrak{g}$; but then two positive roots of $\mathfrak{g}$ will be distinct and proportional, a contradiction. Hence $\mathfrak{k} = \mathfrak{so}(2n, \boldsymbol{C})$ and σ is the natural representation. By an argument similar to that in case A, we see that $\mathfrak{g}_0 = \mathfrak{so}_0(2n, \mathbf{C})$.

Case C. $\mathfrak{k}$ non-semisimple. Let μ_0 be the linear form on $\mathfrak{h}^* = i\mathfrak{h}_{\mathfrak{k}_0}$ defined by $\mu_0(i\mathfrak{h}_{\mathfrak{k}'_0}) = 0$, $\mu_0(H_0) = 1$. From condition 1, we deduce using Lemmas 10 and 12 that if $\mathfrak{k}'$ is the semisimple part of $\mathfrak{k}$, then

$$\mathfrak{k}' = \mathfrak{k}'_1 \times \mathfrak{k}'_2 \cdots \times \mathfrak{k}'_{m-1} \times \mathfrak{k}'_m$$

where $\mathfrak{k}'_i$ $1 \leqq i \leqq m-1$ are algebras of type A_n and $\mathfrak{k}'_m$ is an algebra of one of the types A_n, B_n, C_n or D_n. Let also γ_i be the *unique* simple root of $\mathfrak{k}'_i$ (considered a root of $\mathfrak{k}$) such that $\langle \beta, \gamma_i \rangle_{\mathfrak{k}} > 0$ (Lemma 10). As before this implies that $\langle \beta, \gamma_i \rangle_{\mathfrak{g}} > 0$ (β is the highest weight of σ for which $\beta(H_0) > 0$). It follows then that

$$\beta = \beta(H_0)\mu_0 + \sum_{i=1}^{m} q_i\gamma_i + \sum_{j=1}^{n} r_j\delta_j$$

where $(\delta_j)_{1 \leqq j \leqq n}$ are the simple roots other than the γ_i of $\mathfrak{k}$. Here $\frac{1}{2} \leqq q_i \leqq 1$ (Lemma 10) and $r_j > 0$ (Lemma 7). Now the δ_j and the γ_i are also simple roots of $\mathfrak{g}$ (Lemma 19) so that $\langle \Lambda, \delta_i \rangle_{\mathfrak{g}} \geqq 0$, $\langle \Lambda, \gamma_i \rangle_{\mathfrak{g}} > 0$. It follows that

$$\langle \Lambda, \beta \rangle_{\mathfrak{g}} \geqq \beta(H_0) \cdot \Lambda(H_0) \cdot + \sum_{i=1}^{m} q_i \langle \Lambda, \gamma_i \rangle_{\mathfrak{g}}.$$

On the other hand, since $\langle \beta, \gamma_i \rangle_{\mathfrak{g}} > 0$, we see that $\beta - \gamma_i$ is a root of $\mathfrak{g}$ (Lemma 13) and clearly $\beta - \gamma_i \in C$. It follows then from the facts that $\beta(H_0) > 0$, $\Lambda(H_0) > 0$ (Lemma 20) that

$$1 > \sum_{i=1}^{m} q_i$$

and since $\frac{1}{2} \leqq q_i \leqq 1$, it follows that $m = 0$ or $m = 1$. In the former case $\mathfrak{g}_0 = \mathfrak{sl}(2, \boldsymbol{R})$ and in the latter case, $\mathfrak{k}' = \mathfrak{sl}(n, \boldsymbol{C})$ for some $n (q_1 < 1)$; also since $\sigma_1 |_{\mathfrak{k}'}$ is the natural representation it follows that $\mathfrak{k}_0 = \mathfrak{u}(n)$ and σ_0 is the natural representation in $\boldsymbol{R}^{2n} = \boldsymbol{C}^n$. It follows that $\mathfrak{g}_0 = \mathfrak{su}(n, 1)$: we apply an argument similar to the ones given earlier.

That the representations ρ are the corresponding natural representation is not very difficult to prove.

Case **A**. It turns out if we represent the algebra $\mathfrak{g}$ $(=\mathfrak{so}(2n+2,\boldsymbol{C})$, $\mathfrak{k}=\mathfrak{so}(2n+1))$ by the matrices which leave invariant the quadratic form (on $\boldsymbol{C}^{2n+2}$)

$$\begin{pmatrix} 0 & I_{n+1} \\ I_{n+1} & 0 \end{pmatrix}$$

that $\mathfrak{h}$ is the algebra of diagonal matrices of the form

$$\begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix}$$

and $\theta/\mathfrak{h}$ is the involution

$$\theta\left\{\begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix}\right\} = \left\{\begin{matrix} H' & 0 \\ 0 & -H \end{matrix}\right\}$$

where $H=\begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & h_n \end{pmatrix}$ and $H'=\begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & -h_n \end{pmatrix}$ (see exposé 14, [7]).

If λ_i is the linear form

$$\lambda_i \left\{\begin{pmatrix} H & 0 \\ 0 & -H \end{pmatrix}\right\} = h_i,$$

then for the usual ordering of the basis of the above $\mathfrak{h}$ (which is not in violation of our method of choice), that the roots of $\mathfrak{g}$ are $\pm\lambda_i\pm\lambda_i$ $(i\neq j)$ $1\leqq i\leqq n+1$. Further it is easy to see that $\pm\lambda_i\pm\lambda_j\in A$ for $1\leqq i,j<n$ while $\pm\lambda_i\pm\lambda_n\in B(i\neq n)$ and $C=\emptyset$. Clearly then $r(\lambda_1\pm\lambda_n)$ is the highest weight of σ. It follows then that $\langle\lambda,\pm\lambda_i\pm\lambda_n\rangle=0$ for $i\neq 1$ from the assumption (ii) of the lemma. Also we have

$$\langle\lambda,\lambda_i-\lambda_{i+1}\rangle=\langle\lambda,(\lambda_i-\lambda_{n+1})-(\lambda_{i+1}-\lambda_{n+1})\rangle=0$$

if $i\neq 1$. It follows that the highest weight Λ is of the form $m\mu_N$ where μ_N is the highest weight of the natural representation. We omit the proof for cases **B** and **C** since it is analogous to the above.

In the case of $\mathfrak{su}(n,1)$ we must identify $\mathfrak{sl}(n+1,\boldsymbol{C})$ with $\mathfrak{g}$ in such a way that the complementary simple root (Lemma 19) is the smallest of the simple roots in the lexicographic order to get the natural representation.

4. Proof of Theorem 1″. It is easy to see that Theorem 1″ is equivalent to the following

THEOREM. *Let $\mathfrak{g}_0$ be a real simple non-compact Lie algebra and let $\mathfrak{g}=\mathfrak{g}_0\otimes\boldsymbol{C}$. Let ρ be a non-trivial irreducible representation of $\mathfrak{g}$ with*

highest weight Λ *in a complex vector space* V, *then the operator* T_ρ *on* $V\otimes\mathfrak{p}$ *defined by*

$$2T_\rho = 2(\rho\otimes 1)(c) + (1\otimes\sigma)(c') - (\rho\otimes 1)(c') - (\rho\otimes\sigma)(c')$$

(*which is hermitian symmetric for a suitable scalar product on* $V\otimes\mathfrak{p}$) *is non-negative positive definite in all except possibly the following cases.*

(i) $\mathfrak{g}\cong\mathfrak{so}_0(n,1)$ *in such a way that through this isomorphism,* $\Lambda = m\cdot\mu_N$.

(ii) $\mathfrak{g}\cong\mathfrak{su}(n,1)$ *in such a way that through this isomorphism,* $\Lambda = m\cdot\mu_N$.

We obtain the eigenvalue of T_ρ as follows. Let $V=\sum_{t=1}^{N} V_t$ be a direct sum decomposition into irreducible components of V with respect to the algebra $\mathfrak{k}$: $\rho\,|_{\mathfrak{k}}$, it is well known, is completely reducible. We denote by $\rho(t)$ the corresponding representation of $\mathfrak{k}$ in V_t. Similarly let $\mathfrak{p}=\mathfrak{p}_1\oplus\mathfrak{p}_2$ and $\sigma=\sigma_1\oplus\sigma_2$: here we set $\mathfrak{p}_2=0$ when σ, the adjoint representation of $\mathfrak{k}$ in $\mathfrak{p}$ is irreducible. Further σ_1 is that component of σ such that $\sigma_1(H)$ is a positive scalar. Clearly, each of the subspaces $V_t\underset{C}{\otimes}\mathfrak{p}_k$ is stable under $\mathfrak{k}$. Let then $V_t\underset{C}{\otimes}\mathfrak{p}_k=\sum_l W(t,k,l)$ (direct sum) and correspondingly $\rho(t)\otimes\sigma_k=\rho(t,k)$ $=\sum_l\rho(t,k,l)$. Now in the expression for the operator T_ρ, each of the operators $2(\rho\otimes 1)(c)$ and $(1\otimes\sigma)(c')$ are scalars on the whole $V\underset{C}{\otimes}\mathfrak{p}$ while $(\rho\otimes 1)(c')$ is a scalar on each of the spaces $V_t\underset{C}{\otimes}\mathfrak{p}_k$. Finally, $(\rho\otimes\sigma)(c')$ is a scalar operator on each $W(t,k,l)$. It follows that each $W(t,k,l)$ is contained in an eigen-subspace of T_ρ. We denote the eigen value of T_ρ in $W(t,k,l)$ by $a(t,k,l)$. Also, we denote by $\Lambda(t,k,l)$ the highest weight of $\rho(t,k,l)$ and assume the indexing by l, so chosen, that, $\Lambda(t,k,1)=\Lambda_t+\beta_k$ where Λ_t is the highest weight of ρ_t and β_k that of σ_k. In the sequel, we mean by $\rho(t,k,l)(c')$ also the scalar as well as the operator.

Assertion I. $a(t,k,l)\geqq a(t,k,1)$.

Proof. Since

$$2T_\rho = 2(\rho\otimes 1)(c) + (1\otimes\sigma)(c') - (\rho\otimes 1)(c') - (\rho\otimes\sigma)(c')$$

and the first three operators are scalar operators on the whole of $V_t\underset{C}{\otimes}\mathfrak{p}$, it is sufficient to prove that

$$\rho(t,k,1)(c')\geqq\rho(t,k,l)(c').$$

Now since $\rho(t,k,l)$ is a component of $\rho_t \otimes \sigma_k$, we have

$$\Lambda(t,k,l) = \Lambda_t + \beta_k - \sum_{i=1}^{p} m_i \gamma_i$$

where $(\gamma_i)_{1 \leq i \leq p}$ are the simple roots of $\mathfrak{k}$ and $m_i \geqq 0$. We have further $c' = c'_1 + c'_2$ where $c'_1 = H_0{}^2$ and c'_2 is the "Casimir element" of $B_{\mathfrak{g}}$ restricted to $\mathfrak{k}'$, the semisimple part of $\mathfrak{k}$. Now $\gamma_i(H_0) = 0$ so that we have

$$\rho(t,k,l)(c'_1) = \Lambda(t,k,l)(H_0)^2 = (\Lambda_t + \beta_k)(H_0)^2 = \rho(t,k,1)(c'_1)$$

On the other hand, since $B_{\mathfrak{g}}$ is a positive scalar multiple of the Killing form on each simple component of $\mathfrak{k}'$, we have from Lemma 5,

$$\rho(t,k,1)(c'_2) \geqq \rho(t,k,l)(c'_2).$$

It follows that $\rho(t,k,1)(c') \geqq \rho(t,k,l)(c')$ and the assertion follows.

Assertion I reduces the problem to proving that $a(t,k,1) > 0$ for all t. We divide $[1,N]$ into two parts S and S'; $S = \{t \mid 1 \leqq t \leqq N,\ \Lambda_t(H_0) \geqq 0\}$; S' is the complement of S. We order the indexing set as follows: $S = (t_1, t_2, \cdots, t_{N'})$ is so ordered that $\Lambda_{t_i} \geqq \Lambda_{t_{i'}}$, if $t_i \leqq t_{i'}$. We will prove that $a(t_i, k, 1) > 0$ $(t \in S)$ by induction on i; the proof for the $a(t,k,1)$, $t \in S'$ is analogous (we need only change the order on $\mathfrak{h}^*$ by changing H_0 to $-H_0$). In the sequel we denote Λ_{t_i} simply Λ_i and instead of N' write N; also $a(t_i,k,1)$ is denoted $a(i,k,1)$. Let further, $P_\sigma = \{\alpha \mid \alpha \in \Delta^+, r(\alpha)$ restricted to $\mathfrak{h}_{\mathfrak{k}'}$ is a non-negative linear combination of simple roots of $\mathfrak{k}'\}$. ($\mathfrak{k}'$ is the semi-simple part of $\mathfrak{k}$). In the sequel *we assume that rank* $\mathfrak{g} > 1$; when rank $\mathfrak{g} = 1$ every representation falls in (ii); $P_\sigma = \phi$ and an easy direct computation shows that $T_\rho \geqq 0$.

Assertion II. Suppose that $a(t,k,1) \geqq 0$ for $t < t_0$. Let v_{t_0} be the highest weight vector of Λ_{t_0}. Suppose there is an $\alpha \in P_\sigma$ such that $E_\alpha \cdot v_{t_0} \neq 0$, then $a(t_0,k,1) > 0$.

Proof. $T_\rho = 2(\rho \otimes 1)(c) + (1 \otimes \sigma)(c') - (\rho \otimes 1)(c') - (\rho \otimes \sigma)(c')$ and the first two of the operators on the right are scalars on the whole of $V \underset{C}{\otimes} \mathfrak{p}_k$. Now since $E_\alpha v_{t_0} \neq 0$, it must be a weight vector of weight $\Lambda_{t_0} + r(\alpha)$. Clearly, we have $\Lambda_{t_0} + r(\alpha) > \Lambda_{t_0}$. (See (f) of Lemma 16). Hence $\Lambda_{t_0} + r(\alpha)$ is a weight of ρ_t with $t < t_0$. It follows that

$$\Lambda_t - \Lambda_{t_0} = \sum_{i=1}^{p} m_i \gamma_i + r(\alpha) \cdots \qquad \text{(A)}$$

where $\gamma_1,\cdots,\gamma_p$ are the simple roots of $\mathfrak{k}$ and $m_i \geqq 0$; moreover since $\alpha \in P_\sigma$, $r(\alpha)|_{\mathfrak{h}_{\mathfrak{k}'}} = \sum_{i=1}^{p} n_i\gamma_i$ with $n_i \geqq 0$. Now if we write as before $c' = c'_1 + c'_2$, where $c'_1 = H_0{}^2$ and c'_2 is the Casimir element of $B_{\mathfrak{g}}|_{\mathfrak{k}'}$, we have,

$$\begin{aligned}(\rho_t \otimes 1)(c'_1) - (\rho_{t_0} \otimes 1)(c'_1) &= \Lambda_t(H_0)^2 - \Lambda_{t_0}(H_0)^2 \\ &= r(\alpha)(H_0)(2\Lambda_t(H_0) + r(\alpha)(H_0))\end{aligned}$$

since $\Lambda_t(H_0) = \Lambda_{t_0}(H_0) + r(\alpha)(H_0)$. Also we have

$$\begin{aligned}\{\rho(t,k,1) - \rho(t_0,k,1)\}(c'_1) &= (\Lambda_t + \beta_k)(H_0)^2 - (\Lambda_{t_0} + \beta_k)(H_0)^2 \\ &= 2(\Lambda_t + \beta_k)(H_0)r(\alpha)(H_0) + r(\alpha)(H_0)^2.\end{aligned}$$

It follows that

$$\begin{aligned}\rho(t,k,1)(c'_1) + (\rho_t \otimes 1)(c'_1) - \rho(t_0,k,1)(c'_1) - (\rho_{t_0} \otimes 1)(c'_1) \\ = r(\alpha)(H_0)\{4\Lambda_t(H_0) + 2\beta_k(H_0) + 2r(\alpha)(H_0)\}\end{aligned}$$

and since $\beta_k(H_0) = \pm r(\alpha)(H_0)$, the right side is greater than zero. On the other hand using Lemma 5 and the equation (A) above and the similar equation

$$(\beta_k + \Lambda_t) - (\Lambda_{t_0} + \beta_k) = \sum_{i=1}^{p} m_i\gamma_i + r(\alpha)$$

$(\alpha \in P_\sigma$; i.e. $r(\alpha)|_{\mathfrak{h}_{\mathfrak{k}'}} = \sum_{i=1}^{k} n_i\gamma_i,\ n_i \geqq 0)$, we conclude that

$$\{\rho(t,k,1) + (\rho_t \otimes 1) - \rho(t_0,k,1) - (\rho_{t_0} \otimes 1)\}(c'_2) > 0.$$

It follows that

$$\{\rho_t \otimes 1 + \rho(t,k,1) - (\rho_{t_0} \otimes 1) - \rho(t_0,k,1)\}(c') > 0.$$

Since, as remarked earlier, $2(\rho \otimes 1)(c)$ and $(1 \otimes \sigma)(c')$ are scalars independent of t, we find that

$$a(t_0,k,1) > a(t,k,1)$$

where $t < t_0$. It follows from the hypothesis that $a(t\,,k,1) > 0$.

Assertion III. As before, assume $a(t,k,1) > 0$ for $t < t_0$. Suppose for $\alpha \in P_\sigma$, $E_\alpha v_{t_0} = 0$ but that there exists $\alpha_0 \in \Delta^+$ such that $E_{\alpha_0} v_{t_0} \neq 0$; then $a(t_0,k,1) > 0$.

Proof. We have for $\alpha \in P_\sigma$,

$$\text{(I)} \qquad \cdots (E_\alpha E_{-\alpha} + E_{-\alpha}E_\alpha)(v_{t_0}) = (E_\alpha E_{-\alpha} - E_{-\alpha}E_\alpha)(v_{t_0}) = H_\alpha \cdot v_{t_0}.$$

(we have omitted writing ρ). We have now

$$\rho(c)\cdot v_{t_0}=\langle\rho(c)(v_{t_0}),v_{t_0}\rangle\cdot v_{t_0}$$

where $\langle\ ,\ \rangle$ denotes a hermitian scalar product on V such that $\rho(X)$ for $X\in\mathfrak{k}_0$ is skew-hermitian and $\rho(X)$ for $X\in\mathfrak{p}$ is hermitian symmetric: with respect to such a scalar product, the restriction of ρ to any of the three dimensional algebras $\boldsymbol{C}E_\alpha\oplus\boldsymbol{C}E_{-\alpha}\oplus\boldsymbol{C}H_\alpha$, $\alpha\in\Delta^+$, breaks up into mutually orthogonal irreducible components. We have moreover

$$E_\alpha v_t=0 \text{ for } \alpha\in A^+\cup B^+$$

(If $\alpha\in A^+\cup B^+$, clearly $r(\alpha)$ is a non-negative linear combination of simple roots of $\mathfrak{k}$). Hence we obtain, using (I) above that if $\langle v_{t_0},v_{t_0}\rangle=1$,

$$\langle\rho(c)v_{t_0},v_{t_0}\rangle=\sum_{\alpha\in B^+\cup A^+}\Lambda_{t_0}(H_\alpha)+\sum_{i=0}^{p}\Lambda_{t_0}(H_i)^2+\langle\{\rho(\sum_{\alpha\in C^+}E_\alpha E_{-\alpha}+E_\alpha E_{-\alpha})\cdot\\+\rho(\sum_{i=p+1}^{l}H_i^2)\}(v_{t_0}),v_{t_0}\rangle.$$

We note that $H_i\in i\mathfrak{k}_0\oplus\mathfrak{p}_0$ so that $\rho(H_i)^2$ are hermitian symmetric hence $\langle\rho(H_i)^2v_{t_0},v_{t_0}\rangle\geqq 0$. On the other hand combining Lemma 4 (applied to $A=\langle\ ,\ \rangle_{\mathfrak{g}}$ restricted to $\mathfrak{k}$) and Lemma 16(d), we see that

$$(\rho_{t_0}\otimes 1)(c')=\sum_{\alpha\in A^+}\Lambda_{t_0}(H_\alpha)+\tfrac{1}{2}\sum_{\alpha\in B^+}\Lambda_{t_0}(H_\alpha+H_{\alpha^\theta})+\sum_{i=0}^{l}\Lambda_{t_0}(H_i)^2$$

$$(1\otimes\sigma_k)(c')=\sum_{\alpha\in A^+}\beta_k\ H_\alpha)+\sum_{\alpha\in B^+}\beta_k\ \frac{H_\alpha+H_{\alpha^\theta}}{2}+\sum_{i=0}^{l}\beta_k(H_i)^2$$

$$\rho(t_0,k,1)(c')=\sum_{\alpha\in A^+}(\Lambda_{t_0}+\beta_k)(H_\alpha)+\tfrac{1}{2}\sum_{\alpha\in B^+}(\Lambda_{t_0}+\beta_k)(H_{\alpha^\theta})+\sum_{i=0}^{l}(\Lambda_{t_0}+\beta_k)(H_i)^2.$$

It follows that

$$a(t_0,k,1)=2\sum_{\alpha\in B^+}\Lambda_{t_0}\frac{H_\alpha+H_{\alpha^\theta}}{2}\\+\langle 2\rho(\sum_{\alpha\in C^+}E_\alpha E_{-\alpha}+E_{-\alpha}E_\alpha+\sum_{i=p+1}^{l}H_i^2)(v_{t_0}),v_{t_0}\rangle\\-2\sum_{i=0}^{p}\Lambda_{t_0}(H_i)\beta_k(H_i).$$

Case A. σ is irreducible. Then we need only consider the case when $k=1$. It is then clear that $\beta_1=r(\alpha_1)$ for some $\alpha_1 B^+\cup C^+$. If $\alpha_1\in B^+$,

$$\sum_{i=0}^{p}\Lambda_{t_0}(H_i)\beta_k(H_i)=\Lambda_{t_0}\left(\frac{H_{\alpha_1}+H_{\alpha_1^\theta}}{2}\right).$$

It follows then that

$$a(t_0,k,1)=\sum_{\substack{\alpha\in B^+\\ \alpha\neq\alpha_1}}\Lambda_{t_0}\left(\frac{H_\alpha+H_{\alpha^\theta}}{2}\right)$$
$$+\langle 2\rho(\sum_{\alpha\in C^+}E_\alpha E_{-\alpha}+E_{-\alpha}E_\alpha+\sum_{i=p+1}^{l}H_i^2)v_{t_0},v_{t_0}\rangle.$$

The second terms on the right is strictly positive since $\alpha_0\in C^+$ (see Lemma 11; also note that $(H_i)_{p<i\leq l}$ belong to $\mathfrak{p}_0$). The first term is non-negative since Λ_{t_0} is the highest weight of a representation of $\mathfrak{k}$ and $\dfrac{\alpha+\alpha^\theta}{2}$ $(\alpha\in B^+)$ are positive roots of $\mathfrak{k}$. Hence, if $\alpha_1\in B^+$, $a(t,k,1)>0$. Now assume that $\alpha_1\in C^+$,

$$\sum_{i=0}^{p}\Lambda_{t_0}(H_i)\beta_k(H_i)=\Lambda_{t_0}(H_{\alpha_1}).$$

On the other hand, by Lemma 11, we have

$$\langle(E_{\alpha_1}E_{-\alpha_1}+E_{-\alpha_1}E_{\alpha_1})(v_{t_0}),v_{t_0}\rangle\geqq|\Lambda_{t_0}(H_{\alpha_1})|.$$

By the same lemma we have

$$\langle\rho(E_\alpha E_{-\alpha}+E_{-\alpha}E_\alpha)(v_{t_0}),v_{t_0}\rangle\geqq 0\ \text{ for }\ \alpha\in C^+$$

strict inequality holding for $\alpha=\alpha_0$; we have further assumed that $\alpha_0\notin P_\sigma$, hence $\alpha_0\neq\alpha_1$ so that, we have

$$a(t_0,k,1)=2\sum_{\alpha\in B^+}\Lambda_{t_0}\left(\frac{H_\alpha+H_{\alpha^\theta}}{2}\right)+\lambda$$

with $\lambda>0$. Clearly $\Lambda_{t_0}(H_\alpha+H_{\alpha^\theta})\geqq 0$ for every $\alpha\in B^+$. Hence $a(t_0,k,1)>0$. This covers case A completely.

Case B. σ reducible. We have now $\sigma=\sigma_1\oplus\sigma_2$; we recall that if β_k is the highest weight of σ_k, $\beta_1(H_0)=-\beta_2(H_0)>0$. Also, we have rank $\mathfrak{g}$ $=$ rank $\mathfrak{k}$ and the weights of σ are roots of $\mathfrak{g}$ and $B^+=\emptyset$. Hence

$$a(t_0,k,1)=\langle 2\rho\{\sum_{\alpha\in C^+}(E_\alpha E_{-\alpha}+E_{-\alpha}E_\alpha)\}(v_{t_0}),v_{t_0}\rangle-2\sum_{i=0}^{p}\Lambda_{t_0}(H_i)\beta_k(H_i)$$

Now,

$$\sum_{i=0}^{p}\Lambda_{t_0}(H_i)\beta_k(H_i)=\Lambda_{t_0}(H_{\beta_k})$$

Moreover, we have by Lemma 11,

$$\langle\rho(E_{\beta_k}E_{-\beta_k}+E_{-\beta_k}E_{\beta_k})(v_{t_0}),v_{t_0}\rangle\geqq|\Lambda_{t_0}(H_{\beta_k})|.$$

It follows that

$$a(t_0,1,1)\geqq\langle 2\rho\{\sum_{\alpha\in C^+;\,\alpha\neq\beta_1}(E_\alpha E_{-\alpha}+E_{-\alpha}E_\alpha)\}(v_{t_0}),v_{t_0}\rangle$$

Now for $\alpha \in \Delta^+$, $\langle \rho(E_\alpha E_{-\alpha} + E_{-\alpha} E_\alpha)(v_{t_0}), v_{t_0}\rangle \geqq 0$ and

$$\langle \rho(E_{\alpha_0} E_{-\alpha_0} + E_{-\alpha_0} E_{\alpha_0})(v_{t_0}), v_{t_0}\rangle > 0.$$

(see Lemma 11). Clearly $\alpha_0 \in C^+ - \{\beta_1\}$. Hence $a(t_0, 1, 1) > 0$.

Consider now $a(t_0, 2, 1)$. By Lemma 19, the simple roots of $\mathfrak{g}$ are $\alpha_1, \cdots, \alpha_{l-1}, \alpha_l$ where $\alpha_1, \cdots, \alpha_{l-1}$ are simple roots of $\mathfrak{k}$ and $\alpha_l = -\beta_2$. Hence if $E_{-\beta_2} v_{t_0} = 0$, since already $E_{\alpha_i} v_{t_0} = 0$ for $i < l$, $E_\alpha v_{t_0} = 0$ for every $\alpha \in \Delta^+$. We conclude therefore that $E_{-\beta_2} v_{t_0} \neq 0$. Moreover, by Lemma 11,

$$\langle \rho(E_{\beta_2} E_{-\beta_2} + E_{-\beta_2} E_{\beta_2})(v_{t_0}), v_{t_0}\rangle \geqq |\Lambda_{t_0}(H_{\beta_2})|,$$

where strict inequality holds if $\Lambda_{t_0}(H_{\beta_2}) = 0$. It follows that

$$a(t_0, 2, 1) \geqq \langle \rho\{ \sum_{\alpha \in C^+, -\alpha \neq \beta_2} (E_\alpha E_{-\alpha} + E_{-\alpha} E_\alpha)\}(v_{t_0}), v_{t_0}\rangle \geqq 0$$

strict inequality holding if $\Lambda_{t_0}(H_{\beta_2}) = 0$. Hence, using Lemma 11 again, $a(t_0, 2, 1) > 0$ if $\Lambda_{t_0}(H_{\beta_2}) = 0$.

Assume then that $\Lambda_{t_0}(H_{\beta_2}) \neq 0$, i.e.

$$\sum_{i=0}^{p} \Lambda_{t_0}(H_i) \beta_2(H_i) \neq 0.$$

Hence either $\Lambda_{t_0}(H_0) \neq 0$ or $\Lambda_{t_0}(H_i) \neq 0$ for some $i > 0$, i.e., either $\Lambda_{t_0}(H_0) > 0$ or ρ_{t_0} restricted to $\mathfrak{k}'$, is a non-trivial representation of $\mathfrak{k}'$. It follows immediately from Lemma 6 and the fact that $B_{\mathfrak{g}}\,|_{\mathfrak{k}'}$ is a positive scalar multiple of the Killing form of $\mathfrak{k}'$ on each simple component of $\mathfrak{k}'$, that in either case,

$$\sum_{i=0}^{p} \Lambda_{t_0}(H_i) \beta_1(H_i) > 0.$$

i.e.,

$$\Lambda_{t_0}(H_{\beta_1}) > 0.$$

Now $\beta_1 \in C^+$ and $\beta_1 \neq -\beta_2$ and since $E_{\beta_1} v_{t_0} = 0$,

$$\langle \rho(E_{\beta_1} E_{-\beta_1} + E_{-\beta_1} E_{\beta_1})(v_{t_0}), v_{t_0}\rangle = \Lambda_{t_0}(H_{\beta_1}) > 0.$$

Hence using Lemma 11, we conclude again that

$$a(t_0, 2, 1) > 0$$

Assertion IV. v_1 is the highest weight of ρ; hence if $t \neq 1$, $(1 \leqq t \leqq N)$, $E_\alpha v_t \neq 0$ for some $\alpha \in \Delta^+$ (necessarily in $B^+ \cup C^+$).

Proof. We have to show that $E_\alpha v_1 = 0$ for every $\alpha \in \Delta^+$ (this is sufficient, since in an irreducible representation the eigen space corresponding to the highest weight is 1-dimensional). Suppose, in fact, that $E_\alpha v_1 \neq 0$ for some $\alpha \in \Delta^+$; then $\Lambda_1 + r(\alpha)$ is a weight of $\rho\,|_{\mathfrak{k}}$. By Lemma 16, (f), $r(\alpha) > 0$. Hence $\Lambda_1 + r(\alpha) > \Lambda_1$. On the other hand $\Lambda_1 + r(\alpha) \leqq \Lambda_t$ for some t, a contradiction.

Assertions II and III (note that rank $\mathfrak{g} > 1$) reduce the proof of Theorem 1″ to proving

Assertion V. $a(1,k,1) > 0$ if the pairs $(\mathfrak{g}_0, \Lambda\rho_0)$ are all different from $(so_0(n,1), m\cdot\mu_N)$, $(su(n,1), m\cdot\mu_N)$.

Proof. Since v_1 is the highest weight of $\rho, \Lambda_1 = r(\Lambda)$ where Λ is the highest weight of ρ. We have then

(i) $(\rho\otimes 1)(c)(v_1\otimes w_k) = \{\sum_{\alpha\in\Delta^+}\Lambda(H_\alpha) + \sum_{i=0}^{l}\Lambda(H_i)^2\}(v_1\otimes w_k)$

(ii) $(\rho\otimes 1)(c')(v_1\otimes w_k)$

$$= \{\sum_{\alpha\in A}\Lambda(H_\alpha) + \tfrac{1}{2}\sum_{\substack{\alpha\in B^+\\ \alpha>\alpha^\theta}}\Lambda(H_\alpha + H_{\alpha^\theta}) + \sum_{i=0}^{p}\Lambda(H_i)^2\}(v_1\otimes w_k)$$

(iii) $(\rho\otimes\sigma)(c')(v_1\otimes w_k)$

$$= \{\sum_{\alpha\in A^+}(\Lambda+\beta_k)(H_\alpha)$$

$$+ \tfrac{1}{2}\sum_{\substack{\alpha\in B^+\\ \alpha>\alpha^\theta}}(\Lambda+\beta_k)(H_\alpha + H_{\alpha^\theta}) + \sum_{i=0}^{p}(\Lambda+\beta_k)(H_i)^2\}(v_1\otimes w_k)$$

(iv) $(\rho\otimes\sigma)(c')(v\otimes w_k)$

$$= \{\sum_{\alpha\in A^+}\beta_k(H_\alpha) + \tfrac{1}{2}\sum_{\substack{\alpha\in B^+\\ \alpha>\alpha^\theta}}\beta_k(H_\alpha + H_{\alpha^\theta}) + \sum_{i=0}^{p}\beta_k(H_i)^2\}(v_1\otimes w_k).$$

Hence using the fact that, for $\alpha\in\Delta^+$,

$$\sum_{i=0}^{p}\Lambda(H_i)\alpha(H_i) = \tfrac{1}{2}\Lambda(H_\alpha + H_{\alpha^\theta})$$

and that for $\alpha\in A^+\cup C^+$, $\alpha = \alpha^\theta$, we have

$$a(1,k,1) = \sum_{\substack{\alpha\in B^+\cup C^+\\ \alpha>\alpha^\theta}}\Lambda(H_\alpha + H_{\alpha^\theta}) + 2\sum_{p+1}^{l}\Lambda(H_i^2) - \Lambda(H_{\delta_k} + H_{\delta_k^\theta}).$$

where $\delta_k\in\Delta$ is so chosen that $r(\delta_k) = \beta_k$ and $\delta_k \geqq \delta_k^\theta$.

Now σ_0 is a faithful representation and as a real representation of $\mathfrak{k}_0$, σ_1 is equivalent to σ_0. Now it is easy to see that every ideal (over $\boldsymbol{C}$) of $\mathfrak{k}$ intersects $\mathfrak{k}_0$ non-trivially so that σ_1 is a faithful representation of $\mathfrak{k}$. (The assertion above is a consequence of the fact that the complexification of a *compact* simple algebra is simple.)

On the other hand since ρ is non-trivial, there is an ideal $\mathfrak{g}'$ in $\mathfrak{g}$ such that $\rho\,|_{\mathfrak{g}'}$ is faithful. It follows that

$$\Lambda = \sum_{i=1}^{l} m_i \alpha_i$$

where $\alpha_1, \cdots, \alpha_l$ are the simple roots of $\mathfrak{g}$, $m_i \geqq 0$ and $m_i. > 0$ whenever α_i is a root of $\mathfrak{g}'$. Now by (f), Lemma 6, $r(\alpha_i) > 0$ so that $r(\Lambda)$ is a non-trivial highest weight of $\mathfrak{k}'$. We assert now that

$$\Lambda(H_{\delta_1} + H_{\delta_1^\theta}) > 0.$$

In fact

$$\Lambda(H_{\delta_1} + H_{\delta_1^\theta}) = \sum_{i=0}^{p} \Lambda(H_i)\cdot\delta_1(H_i)\,;$$

Since $r(\Lambda)$ restricted to $\mathfrak{h}_{\mathfrak{k}'}$ is the highest weight of a representation of $\mathfrak{k}'$,

$$\sum_{i=1}^{p} \Lambda(H_i)\cdot\delta_1(H_i) \geqq 0 \quad \text{(Lemma 6, Corollary)}$$

here strict inequality occurs if $\mathfrak{k}' = \mathfrak{k}$; on the other hand if $\mathfrak{k}' \neq \mathfrak{k}$, $H_0 \neq 0$ and $\Lambda(H_0)\delta_1(H_0) > 0$ (see Lemma 20). Thus, in any case

$$\Lambda(H_{\delta_1} + H_{\delta_1^\theta}) > 0$$

Now we have $\Lambda(H_{\delta_2} + H_{\delta_2^\theta}) > 0$ since $\delta_2 = \delta_2^\theta < 0$ (if $\delta_2 \neq 0$) and Λ is the highest weight of ρ. Hence

$$a(1,2,1) \geqq \sum_{\alpha \in B^+ \cup C^+} \Lambda(H_\alpha + H_{\alpha^\theta})$$

Since $\delta_1 \in B^+ \cup C^+$, we conclude that $a(1,2,1) > 0$ ($\Lambda(H_\alpha) \geqq 0$ for $\alpha \in \Delta^+$).

Consider now $a(1,1,1)$. We first assume that $\mathfrak{k}$ is semisimple. Let us denote by $A(x,y)$ the restriction of $B_{\mathfrak{g}}$ to $\mathfrak{k}$ as also the induced scalar product on the dual of $\mathfrak{h}_{\mathfrak{k}}$. Then for linear forms λ_1, λ_2 on $\mathfrak{h}$,

$$A(r(\lambda_1), r(\lambda_2)) = \sum_{i=1}^{p} \lambda_1(H_i)\lambda_2(H_i).$$

Adopting then the notations of Lemma 10 (σ now playing the role of ρ of that lemma), we have

$$a(1,1,1) = 2A(r(\Lambda),\mu_\sigma) + \sum_{B^+\cup C^+;\,\alpha\geqq\alpha^\theta;\,\alpha\notin P\sigma} \Lambda(H_\alpha + H_{\alpha^\theta}) + 2\sum_{i=p+1}^{l} \Lambda(H_i)^2.$$

Since $r(\Lambda)\neq 0$, $A(r(\Lambda),\mu_\sigma) > 0$ if σ is of type P. Hence if σ is of type P, $a(1,1,1) > 0$. Even if σ is not of type P, $a(1,1,1) \geqq 0$ and $= 0$ only if $\Lambda(H_\alpha) = 0$ for every $\alpha \in B^+ \cup C^+$ such that $\alpha \neq \delta_1$ or δ_1^θ. By Lemma 23, it follows that $\mathfrak{g}_0 = \mathfrak{so}(n,1)$ and that $\Lambda = m\cdot\mu_N$.

Assume now that $\mathfrak{k}$ is not semisimple. We denote now by $A(x,y)$ the restriction of $B_\mathfrak{g}$ to $\mathfrak{k}'$ as also the induced scalar product on the dual of $\mathfrak{h}_{\mathfrak{k}'}$. Let $r_1(\lambda)$ denote the restriction of a linear form λ on $\mathfrak{h}$ to $\mathfrak{h}'_{\mathfrak{k}}$. Then we have

$$a(1,1,1) = 2A(r_1(\Lambda),\mu_{\sigma_1'}) + 2\sum_{\substack{\alpha\in C^+\\ \alpha\notin P\sigma}} \Lambda(H_\alpha) + \sum_{\substack{\alpha\in P\sigma\\ \alpha\neq\delta_1}} \Lambda(H_0)\cdot\alpha(H_0).$$

Now each of the terms on the right are non-negative. Hence $a(1,1,1)\geqq 0$ and $= 0$ only if σ_1 is not of type P and $\Lambda(H_\alpha) = 0$ for every $\alpha\in C^+$ different from δ_1. Applying Lemma 23, we conclude that $\mathfrak{g}_0 = \mathfrak{su}(n,1)$, $\mathfrak{g} = \mathfrak{ls}(n+1,C)$ and $\Lambda = m\cdot\mu_N$.

This completes the proof of Theorem 1″.

Remark. Completion of the proof of the corollary. We need only check that when $\mathfrak{g}_0$ is one of the algebras $\mathfrak{so}_0(n,1)$ or $\mathfrak{su}(n,1)$ and Λ the highest weight of the adjoint representation of $\mathfrak{g} = \mathfrak{so}(n+1,\boldsymbol{C})$ or $\mathfrak{sl}(n+1,\boldsymbol{C})$, $\Lambda \neq m\mu_N$ unless $n = 2$ or $n = 1$ respectively. Let $\mathfrak{g}_0 = \mathfrak{so}_0(2n,1)$ ($\mathfrak{g} = \mathfrak{so}(2n+1,\boldsymbol{C})$ for instance); if we take for $\mathfrak{h}$ the standard Cartan subalgebra (see exposé 14, [7]) the highest root is $e_1 + e_2$ whereas the highest weight of the natural representation is e_1. Hence except when $\mathfrak{g} = \mathfrak{sl}(2,\boldsymbol{C}) = \mathfrak{so}(3,\boldsymbol{C})$, $e_1 + e_2 \neq me_1$. Hence the corollary (the proof for $\mathfrak{g}_0 = \mathfrak{su}(n,1)$ etc. is similar.)

TATA INSTITUTE OF FUNDAMENTAL RESEARCH,
BOMBAY.

REFERENCES.

[1] N. Bourbaki, *Éléments de mathématiques, Groupes et algèbres de Lie, Chapitre 1, Algèbres de Lie*, Hermann, Paris.

[2] S. Eilenberg, "Homology of spaces with operators, I," *Transactions of the American Mathematical Society*, vol. 61 (1947), pp. 378-417.

[3] S. Helgason, *Differential geometry and symmetric spaces*, Academic Press, 1962.

[4] Y. Matsushima, "On the first Betti number of compact quotient spaces of higher dimensional symmetric spaces," *Annals of Mathematics*, vol. 75 (1962), pp. 312-330.

[5] Y. Matsushima and S. Murakami, "On vector bundle valued harmonic forms and automorphic forms on symmetric Riemannian manifolds," *Annals of Mathematics*, vol. 78 (1963), pp. 365-416.

[6] M. S. Raghunathan, "Deformations of linear connections and Riemannian manifolds," *Journal of Mathematics and Mechanics*, vol. 13 (1964), pp. 97-123.

[7] Séminaire "Sophus Lie," 1 e année: 1954/1955, *Théorie des algèbre de Lie, Topologie des groupes de Lie, Secrétariat mathematique*, Paris, 1955.

[8] A. Weil, "Discrete subgroups of Lie groups, II," *Annals of Mathematics*, vol. 75 1962), pp. 97-123.

LOCAL DIFFEOMORPHISMS—C^∞ REALIZATION OF FORMAL PROPERTIES.

By Kuo-Tsai Chen.[1]

Let R^n be the real n-space with a given system of coordinates $x = (x^i)$. By a local diffeomorphism we mean one of the class C^∞ which maps a neighborhood of the origin O onto another and leaves 0 fixed. The totality of such local diffeomorphisms forms a group G. Here, of course, as elements of the group G, local diffeomorphisms are taken in the sense of their germs, i.e., two local diffeomorphisms are considered to be the same element of the group if they coincide in some neighborhood of 0.

Denote by G_∞ the normal subgroups of G which consists of all local diffeomorphisms having contact of ∞ order with the identity map of R^n. The quotient group $\mathfrak{G} = G/G_\infty$ will be called the group of formal transformations. Let $\otimes : G \to \mathfrak{G}$ be the natural homomorphism. Each formal transformation $\otimes T$ can also be considered as the substitution of the formal power series algebra $\mathfrak{F}$ in x over R such that each x^i is substituted by the Taylor's expansion of $x^i \circ T$.

We shall deal with the following notions: (a) equivalence, (b) semisimple part, (c) logarithm.

The equivalence between elements of G (or $\mathfrak{G}$) is in the sense of conjugation in the group.

Recall that a linear transformation of a vector space V is said to be semisimple if any invariant subspace of V has an invariant complement. For the case of V being finite dimensional over the complex number field C, a linear transformation is semisimple if and only if it can be represented by a diagonal matrix. Also recall that for any $T \in G$ (or $\hat{T} \in \mathfrak{G}$), the jacobian $J(T)$ (or $J(\hat{T})$) at 0 is a nonsingular linear transformation of R^n.

We define an element of G (or $\mathfrak{G}$) to be semisimple if it is equivalent to a semisimple linear transformation with respect to the given coordinates x. A semisimple part ϕ of $T \in G$ is an element of G such that

(a) ϕ is semisimple in G,

Received May 7, 1964.

[1] The work has been partially supported by the National Science Foundation under Grant NSF-GP-2563.

(b) ϕ and T commute,

(c) $J(\phi^{-1}T) - E$ is nilpotent,

where E is the identity map of R^n. The semisimple part of any element of $\mathfrak{G}$ can be defined precisely in the same manner.

For any vector field X, denote by $\operatorname{Exp} tX\, p$ (instead of $\exp tX\, p$ as in [2]) the intergral path of X that passes through the point p when $t = 0$. If X is of C^∞ over a neighborhood of 0 and vanishes at 0, then $\operatorname{Exp} X$ is a local diffeomorphism. If $T \in G$ is such a local diffeomorphism, then we say that X is a logarithm of T. Logarithms of a formal transformation can be similarly defined to be formal vector fields. (see [2])

In § 1, we show that any formal transformation has a unique semisimple part. In § 4, we establish a necessary and sufficient condition (Theorem 4. 1) for any formal transformation to possess a logarithm. Consequently we give a new proof of a result (Theorem 4. 2) due to D. C. Lewis [6], which states that some power of a formal transformation possesses a logarithm.

A local diffeomorphism T is said to have an elementary fixed point at 0 if $J(T)$ has no eigen value lying on the unit circle of the complex plane. Our results in the direction of C^∞ realization of formal properties hold for such local diffeomorphisms and can be summarized in the next three theorems, of which Theorem 2 is due to S. Sternberg [9].

THEOREM 1. *If 0 is an elementary fixed point of $T \in G$, then the semisimple part ϕ of T exists such that $\circledS\phi$ is the semisimple part of $\circledS T$. Moreover, ψ is another semisimple part of T if and only if $\phi^{-1}\psi$ belongs to G_∞ and commutes with T.*

THEOREM 2. *Let 0 be an elementary fixed point of T, $U \in G$. Then T and U are equivalent in G if and only if $\circledS T$ and $\circledS U$ are equivalent in $\mathfrak{G}$.*

THEOREM 3. *Let 0 be an elementary fixed point of $T \in G$. Then T possesses a logarithm. if and only if so does $\circledS T$. Moreover, there always exists a positive integer α depending only on $J(T)$ such that T^α possesses a logarithm.*

Finally, to illustrate that α may take any positive integral value, we construct, for any preassigned α, a local diffeomorphism T in R^3 such that T^β possesses a logarithm if and only if

$$\beta = 0 \bmod \alpha.$$

For the proof of Theorem 1, assume without loss of generality that the

semisimple part $\hat{\phi}$ of $\circledS T$ is already linear so that $\hat{\phi}$ is also well defined as a C^∞ diffeomorphism. We first construct a local diffeomorphism T' having $\hat{\phi}$ as a semisimple part such that $\circledS T = \circledS T'$ (see § 2). Then T and T' are shown to be equivalent via a modification of Sternberg's wedge method (Lemma 3.1, [2]). This leads to the existence of a semisimple part ϕ of T. The remaining portion of the theorem follows from the uniqueness of the semisimple part of $\circledS T$.

Since the construction of T' depends only on $\circledS T$, Theorem 2 is proved simultaneously with Theorem 1. The existence of the stable manifold then follows as a corollary. Sternberg's proof of Theorem 2 depends on the stable manifold, for which he gave a brief proof under the illustrating assumption that the jacobian of the local diffeomorphism in question is already in diagonal form [8].

Theorem 3 follows from Theorem 2 and Theorem 4.2.

The question of C^1-equivalence of local diffeomorphisms has been investigated by P. Hartman [4]. We also wish to mention Sternberg's results [9] on volume preserving diffeomorphisms and their logarithms.

1. Formal simple part. As in § 7 of [2], denote by F the algebra of the germs of C^∞ functions about 0, by A the Lie algebra of the germs of C^∞ vector fields about the critical point 0 and by G the group of the germs of C^∞ diffeomorphisms about the fixed point 0. Denote respectively by $\mathfrak{F}$, $\mathfrak{A}$ and $\mathfrak{G}$ the algebra of formal power series in x, the Lie algebra of formal vector fields and the group of formal transformations. We recall the homeomorphisms $\circledS : F \to \mathfrak{F}$, $\circledS : A \to \mathfrak{A}$ and $\circledS : G \to \mathfrak{G}$, which correspond to taking formal power series expansions. Moreover, $\mathfrak{F}$ and $\mathfrak{A}$ have graded structures through the projections $\pi_r : \mathfrak{F} \to \mathfrak{F}^{(r)}$ and $\pi_r : \mathfrak{A} \to \mathfrak{A}^{(r)}$, $r = 0, 1, \cdots$. (Such graded structures are not invariant under C^∞ change of coordinates.)

For any $\sigma \in G$, denote by $\sigma^* : F \to F$ the induced automorphism on F. Throughout this paper σ, ϕ, ψ, η, T, U will denote elements of G, while X, Y, Z, N will denote those of A.

Write

$$\exp X = (\operatorname{Exp} X)^*.$$

For $\hat{X} \in \mathfrak{A}$ with $\hat{X} = \circledS X$, define

$$\operatorname{Exp} \hat{X} = \circledS \operatorname{Exp} X$$

and

$$\exp \hat{X} = (\operatorname{Exp} \hat{X})^*.$$

Since, for any $f, g \in F$, $\exp X(fg) = (\exp Xf)(\exp Xg)$, we have

$$\exp \hat{X}(\hat{f}\hat{g}) = (\exp \hat{X}\hat{f})(\exp \hat{X}\hat{g}).$$

We assert that

$$(d/dt) \exp t\hat{X}\hat{f} = \exp t\hat{X}\,\hat{X}\hat{f}.$$

In fact, if $\hat{X} = \Theta X$ and $\hat{f} = \Theta f$, then

$$\begin{aligned}(d/dt)(\exp t\hat{X}\hat{f}) &= (d/dt)\Theta(\exp tX f)\\ &= \Theta(d/dt)(\exp tX f)\\ &= \Theta(\exp tX\,Xf) = \exp t\hat{X}\,\hat{X}\hat{f}.\end{aligned}$$

PROPOSITION 1.1. *For* $\hat{X} \in \mathfrak{X}$,

$$\exp t\hat{X}\hat{f} = \sum_{r=0}^{\infty} t^r \hat{X}^r \hat{f}/r!.$$

More precisely, for any $s \geqq 0, \sum_{r=0}^{\infty} t^r \pi_s(\hat{X}^r f)/r!$ *converges to* $\pi_s(\exp t\hat{X} f)$ *for all values of* t.

Proof. What we precisely want to show is that the Taylor's expansion of the $\mathfrak{F}^{(s)}$-valued function $\pi_s(\exp t\hat{X}\hat{f})$ in t converges for all values of t, i.e. $\pi_s(\exp t\hat{X}\hat{f})$ is an entire function for $s \geqq 0$. Use induction on s. The case $s = 0$ is trivial. Assume $s \geqq 0$. If $\hat{f}$ is of the form $\hat{f}_1\hat{f}_2$ with $\pi_0\hat{f}_1 = \pi_0\hat{f}_2 = 0$, then

$$\begin{aligned}\pi_{s+1} \exp t\hat{X}(\hat{f}_1\hat{f}_2) &= \pi_{s+1}((\exp t\hat{X}\hat{f}_1)(\exp t\hat{X}\hat{f}_2))\\ &= \sum_{\lambda=1}^{s} (\pi_\lambda \exp t\hat{X}\hat{f}_1)(\pi_{s+1-\lambda} \exp t\hat{X}\hat{f}_2)\\ &= \sum_{\lambda=1}^{s} \Big(\sum_{\mu=0}^{\infty} t^\mu \pi_\lambda(\hat{X}^\mu \hat{f}_1)/\mu!\Big)\Big(\sum_{\nu=0}^{\infty} t^\nu \pi_{s+1-\lambda}(\hat{X}^\nu \hat{f}_2)/\nu!\Big)\\ &= \sum_{r=0}^{\infty} \sum_{\mu+\nu=r} \sum_{\lambda=1}^{s} t^r C_{r\mu} \pi_\lambda(\hat{X}^\mu \hat{f}_1)\pi_{s+1-\lambda}(\hat{X}^\nu \hat{f}_2)/r!\\ &= \sum_{r=0}^{\infty} \sum_{\mu+\nu=r} t^r C_{r\mu} \pi_{s+1}[(\hat{X}^\mu \hat{f}_1)(\hat{X}^\nu \hat{f}_2)]/r!\\ &= \sum_{r=0}^{\infty} t^r \pi_{s+1}[\hat{X}^r(\hat{f}_1\hat{f}_2)]/r!.\end{aligned}$$

It is easy to see that any $\hat{g} \in \mathfrak{F}$ with $\pi_0\hat{g} = \pi_1\hat{g} = 0$ can be expressed as the sum of a finite number of such products $\hat{f}_1\hat{f}_2$. Therefore the case $s+1$ is valid for all $\hat{f}$ with $\pi_0\hat{f} = \pi_1\hat{f} = 0$. It remains to show the case $s+1$ with $\hat{f} = x^i$.

Now write $h^i = \pi_{s+1}(\exp t\hat{\mathfrak{X}} x^i)$ and $\pi_1(\hat{\mathfrak{X}} x^i) = \sum a_j{}^i x^j$. Then

$$(d/dt)h^i = \pi_{s+1}(d/dt)(\exp t\hat{\mathfrak{X}} x^i)$$
$$= \pi_{s+1} \exp t\hat{\mathfrak{X}}(\sum a_j{}^i x^j + \cdots)$$

so that

$$(1.1) \qquad (d/dt)h^i = \sum a_j{}^i h^j + p^i(t),$$

where p^i is an $\mathfrak{F}^{(s+1)}$-valued entire function of t. Since (1.1) gives a system of n nonhomogeneous linear differential equations in the unknowns h^i, we are led to the conclusion that each h^i is an entire function of t. Any $\hat{f} \in \mathfrak{F}$ may be written as

$$\hat{f} = c + \sum c_i x^i + \hat{g}$$

with $\pi_0 \hat{g} = \pi_1 \hat{g} = 0$. Hence the induction is completed.

Recall the notion of Adj such that, for $\sigma \in G$, $X \in A$, $\hat{\sigma} \in \mathfrak{G}$, $\hat{\mathfrak{X}} \in \mathfrak{A}$,

$$\operatorname{Adj} \sigma X = \sigma^{*-1} X \sigma^*,$$
$$\operatorname{Adj} \hat{\sigma}\hat{\mathfrak{X}} = \hat{\sigma}^{*-1}\hat{\mathfrak{X}}\hat{\sigma}^*.$$

For any two vector spaces V and W over the same field, let $L(V, W)$ be the vector space of the linear maps from V into W. For $g, h \in L(V, V)$, write

$$(\operatorname{ad} h) g = gh - hg.$$

Also recall that, if S is the semisimple part of $X \in L(V, V)$ and if V is finite dimensional over R or C, then

$$(1.2) \qquad \operatorname{im} S \subset \operatorname{im} X,$$

where im = image.

Any $\hat{\mathfrak{X}} \in \mathfrak{A}$ and any $\hat{\sigma}^*$, $\hat{\sigma} \in \mathfrak{G}$, may be taken to belong to $L(\mathfrak{F}, \mathfrak{F})$.

LEMMA 1.1. *If* $\hat{\sigma} \in \mathfrak{G}$ *and if* $\hat{Z} \in \mathfrak{A}$ *with* $\pi_0 Z = 0$, *then*

$$(1.3) \qquad (\exp - t\hat{Z})\hat{\sigma}^*(\exp t\hat{Z}) = \hat{\sigma}^* + \cdots + t^r(\operatorname{ad}\hat{Z})^r \sigma^*/r! + \cdots$$

and, in short, $= (\exp t \operatorname{ad} \hat{Z})\hat{\sigma}^*$.

Proof. Since $\pi_0 \hat{Z} = 0$, for any $\hat{f} \in \mathfrak{F}$, $\pi_s(\operatorname{ad} Z)^r \hat{\sigma}^* \hat{f}$ vanishes when $s < r$, and the right hand side of (1.3) is well defined. Now both

$$u = (\exp t\hat{Z})\hat{\sigma}^*(\exp t\hat{Z})$$

and

$$u = (\exp t \operatorname{ad} \hat{Z})\hat{\sigma}^*$$

satisfy the system of differential equations

$$\partial ux^i/\partial t = ((\operatorname{ad}\hat{Z})u)x^i, \qquad i = 1, \cdots, n,$$

with the initial condition $u(0) = \hat{\sigma}^*$. By term-wise comparison, we have $d\pi_1(ux^i)/dt = 0$, i. e. $\pi_1(ux^i) = \pi_1(\sigma^* x^i)$, and for $r > 1$,

$$d\pi_r(ux^i)/dt = \pi_r((\operatorname{ad}Z)u)x^i = \pi_r(uZx^i - Zux^i),$$

which owing to $\pi_0 Z = 0$, is determined by $\pi_1 ux^j, \cdots, \pi_{r-1}ux^j$, $j = 1, \cdots, n$. It follows that each $\pi_r(ux^i)$ is well determined. This means that the solution u is unique. Hence the lemma is proved.

Remark. The above lemma also holds if $\hat{\sigma}^*$ is replaced by $\hat{\mathfrak{X}} \in \mathfrak{A}$. We have

$$(\operatorname{Adj}\operatorname{Exp} t\hat{Z})\hat{\mathfrak{X}} = (\exp t \operatorname{ad}\hat{Z})\hat{\mathfrak{X}}.$$

In fact, the lemma holds for any linear map

$$\hat{\sigma}^*: \mathfrak{F} \to \mathfrak{F},$$

provided it is order preserving, i. e., for any $\hat{f} \in \mathfrak{F}$, $s \geqq 0$,

$$\pi_0 \hat{f} = \cdots = \pi_s \hat{f} = 0$$

implies

$$\pi_0 \hat{\sigma}^* \hat{f} = \cdots = \pi_s \hat{\sigma}^* \hat{f} = 0.$$

Definition 1.1. For any $\hat{\sigma} \in \mathfrak{G}$, define $J(\hat{\sigma}) \in L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(1)})$ such that

$$J(\hat{\sigma})\hat{f} = \pi_1(\hat{\sigma}^* \hat{f}),\ \hat{f} \in \mathfrak{F}^{(1)}.$$

Definition 1.2. Any $\hat{\phi} \in \mathfrak{G}$ is said to be linear if $\hat{\phi}^* \mathfrak{F}^{(1)} \subset \mathfrak{F}^{(1)}$.

For any linear $\hat{\phi} \in \mathfrak{G}$, define $\operatorname{Ad}_r \hat{\phi}$ to be the linear transformation of $L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(r)})$ such that, for $Z \in L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(r)})$, $\operatorname{Ad}_r \hat{\phi} Z = Z\pi_1 \hat{\phi}^* - \pi_r \hat{\phi}^* Z$.

For any $\hat{T} \in \mathfrak{G}$, denote by $J_r(\hat{T})$ the element of $L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(r)})$ which is the restriction of the mapping $\pi_r \hat{T}^*$ on $\mathfrak{F}^{(1)}$. Thus

$$T^* x^i = J_1(\hat{T})x^i + J_2(\hat{T})x^i + \cdots.$$

PROPOSITION 1.2. *Let $\hat{T}$, $\hat{\phi} \in \mathfrak{G}$ such that $\hat{\phi}$ is linear. Then $\hat{T}$ and $\hat{\phi}$ commute if and only if $J_r(\hat{T}) \in \ker \operatorname{Ad}_r \hat{\phi}$, $r = 1, 2, \cdots$.*

Proof. The commutativity of $\hat{T}$ and $\hat{\phi}$ holds if and only if

$$(\hat{T}\hat{\phi} - \hat{\phi}\hat{T})^* x^i = 0,\ i = 1, \cdots, n,$$

i. e.

$$\pi_r(\hat{T}\hat{\phi}-\hat{\phi}\hat{T})^*x^i=(J_r(\hat{T})\pi_1\hat{\phi}^*-\pi_r\hat{\phi}^*J_r(\hat{T}))x^i$$
$$=(\mathrm{Ad}_r\hat{\phi}J_r(\hat{T}))x^i=0.$$

Hence $J_r(\hat{T})\in\ker\mathrm{Ad}_r\hat{\phi}$.

Now we are going to compute $\ker\mathrm{Ad}_r\hat{\phi}$.

Let $\hat{\phi}\in\mathfrak{G}$ be linear such that $J(\hat{\phi})$ is semisimple. Let $\bar{\mathfrak{F}}$, $\bar{\mathfrak{A}}$ be obtained respectively from $\mathfrak{F}$, $\mathfrak{A}$ by extending the ground field R to the field C of the complex numbers. Denote the respective projections by $\pi_r\colon\bar{\mathfrak{F}}\to\bar{\mathfrak{F}}^{(r)}$ and $\pi_r\colon\bar{\mathfrak{A}}\to\bar{\mathfrak{A}}^{(r)}$. For any $\hat{\sigma}\in\mathfrak{G}$, the substitution $\hat{\sigma}^*$ on $\mathfrak{F}$ can be extended to a substitution $\hat{\sigma}^*$ on $\bar{\mathfrak{F}}$. There exists a base $z^1,\cdots,z^n$ of $\bar{\mathfrak{F}}^{(1)}$ with the following properties:

(a) If z^i is complex, then some z^j is its conjugate.

(b) $\bar{\phi}^*z^i=\lambda_iz^i$, $i=1,\cdots,n$.

In the next lemma, we identify $\bar{\mathfrak{A}}^{(r-1)}$ and $L(\bar{\mathfrak{F}}^{(1)},\bar{\mathfrak{F}}^{(r)})$.

LEMMA 1.2. *If $\hat{\phi}\in\mathfrak{G}$ is linear with $J(\hat{\phi})$ being semisimple as given above, then $\mathrm{Ad}_r\hat{\phi}$ is also semisimple, and $\ker\mathrm{Ad}_r\bar{\phi}$ possesses a base consisting of all $z^m\partial/\partial z^i$ with*

$$\lambda^m=\lambda^i,\qquad |m|=r,\ i=1,\cdots,n. \tag{1.4}$$

Proof. For any monomial z^m, we have $\bar{\phi}^*z^m=\lambda^mz^m$. Now $\bar{\mathfrak{A}}^{(r-1)}$ has a base consisting of all $z^m\partial/\partial z^i$, $|m|=r$, $i=1,\cdots,n$. For $\bar{f}\in\bar{\mathfrak{F}}^{(1)}$, we have

$$\mathrm{Ad}_r\bar{\phi}(z^m\partial/\partial z^i))\bar{f}=\lambda^mz^m\partial\bar{f}/\partial z^i-z^m\partial\bar{\phi}^*\bar{f}/\partial z^i=(\lambda^m-\lambda^i)(z^m\partial/\partial z^i)\bar{f}.$$

Hence the lemma is proved.

LEMMA 1.3. *Let $\hat{\phi},\hat{\psi}\in\mathfrak{G}$. If $J(\hat{\phi})$ is the semisimple part of $J(\hat{\psi})$, then $\mathrm{Ad}_r\hat{\phi}$ is the semisimple part of $\mathrm{Ad}_r\hat{\psi}$.*

Proof. Again, we may assume that $\hat{\phi}$ and $\hat{\psi}$ are linear. It is clear that $\mathrm{Ad}_r\hat{\phi}$ and $\mathrm{Ad}_r\hat{\psi}$ commute. Moreover, the last lemma asserts that $\mathrm{Ad}_r\hat{\phi}$ is semisimple. It suffices to show that $\mathrm{Ad}_r\hat{\phi}-\mathrm{Ad}_r\hat{\psi}$ is nilpotent.

Denote by $\bar{\phi}^*$ and $\bar{\psi}^*$ the respective extension of $\hat{\phi}^*$ and $\hat{\psi}^*$. We may further demand that

$$\bar{\psi}^*z^i=\lambda_iz^i+\epsilon_iz^{i+1}$$

where ϵ_i is equal to either 0 or 1, $i=1,\cdots,n$. Let $\bar{\psi}_r$ and $\bar{\phi}_r$ be respectively the restrictions of $\bar{\psi}^*$ and $\bar{\phi}^*$ on $\bar{\mathfrak{F}}^{(r)}$. Write $N_r=\bar{\psi}_r-\bar{\phi}_r$. Then, for $|m|=r$,

$$N_rz^m=\prod(\lambda_iz^i+\epsilon_iz^{i+1})^{m_i}-\prod(\lambda_iz^i)^{m_i}.$$

The totality of z^m, $|m|=r$, forms a base of $\mathfrak{F}^{(r)}$. If we order the base lexicographically, it is not hard to see that $N_r z^m$ is a linear combination of base elements of order lower than z^m. Thus N_r is indeed nilpotent. For $Z \in L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(r)})$,

$$(\mathrm{Ad}_r \psi - \mathrm{Ad}_r \hat{\phi}) Z = N_r Z - Z N_1.$$

It follows that

$$(\mathrm{Ad}_r \psi - \mathrm{Ad}_r \hat{\phi})^l Z = N_r^l Z - C_1^l N_r^{l-1} Z N_1 + \cdots + (-1)^l Z N_1^l$$

which vanishes when l is sufficiently large. Hence the lemma holds.

COROLLARY. *If $\hat{\phi}$ is as given in Lemma 1.2, and if $\hat{T}$ and $\hat{\phi}$ commute, then $\pi_r T^* z^i$ is a linear combination (over C) of the monomials z^m such that $\lambda^m = \lambda^i$, $|m|=r$.*

THEOREM 1.1. *Any $\hat{U} \in \mathfrak{G}$ has a unique semisimple part.*

Proof. Let $\hat{\phi}$ be the (linear) semisimple part of the linear transformation $J(\hat{U})$. For the existence of the semisimple part of $\hat{U}$, it suffices to construct some $\hat{T} \in \mathfrak{G}$ equivalent to $\hat{U}$ and having $\hat{\phi}$ as a semisimple part. Our trick is to seek $\hat{Z} \in \mathfrak{A}$ with $\pi_0 \hat{Z} = 0$ such that

$$\hat{T}^* = (\exp - \hat{Z}) \hat{U} (\exp \hat{Z}).$$

Observe that fact that $\pi_0 \hat{Z} = 0$ implies $J(\hat{T}) = J(\hat{U})$. Thus $J(\hat{\phi}^{-1}\hat{T}) - E$ is indeed nilpotent. The only restriction on $\hat{Z}$ is the commutativity of $\hat{\phi}$ and $\hat{T}$, which holds if and only if

$$J_r(\hat{T}) \in \ker \mathrm{Ad}_r \hat{\phi}. \tag{1.5}$$

Write $\hat{Z}_r = \pi_r \hat{Z}$. We shall determine $\hat{Z}_r$ by induction on r such that (1.5) holds. The case $r=0$ is trivial. For $r>0$ we note that

$$\begin{aligned}\pi_r(\mathrm{ad}\, \hat{Z}\hat{U}^*) x^i = \pi_r(\hat{U}^* \hat{Z} x^i - \hat{Z}\hat{U}^* x^i) &= (\mathrm{Ad}_r \hat{U}\hat{Z}) x^i \\ &+ \text{terms determined by } \hat{U} \text{ and } \hat{Z}_1, \cdots, \hat{Z}_{r-1}.\end{aligned}$$

Thus

$$J_r(\hat{U}) x^i = \pi_r(\exp \mathrm{ad}\, \hat{Z}\hat{U}^*) x^i = (\mathrm{Ad}_r \hat{U} Z_r) x^i + R_r^i$$

where R_r^i is determined by $\hat{U}$ and $\hat{Z}_1, \cdots, \hat{Z}_{r-1}$. This means that $J_r(\hat{U})$ is the sum of $\mathrm{Ad}_r \hat{U} Z_r$ and another element of $L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(r)})$ well determined by the induction hypothesis. Since $\mathrm{Ad}_r \hat{\phi}$ is semisimple,

$$L(\mathfrak{F}^{(1)}, \mathfrak{F}^{(r)}) = \ker \mathrm{Ad}_r \hat{\phi} + \mathrm{im}\, \mathrm{Ad}_r \hat{\phi}.$$

It follows from (1.2) and Lemma 1.3 that we may choose $\hat{Z}_r$ such that

$\mathrm{Ad}_r \hat{U} Z_r$ becomes any preassigned element of $\operatorname{im} \mathrm{Ad}_r \hat{\phi}$. In particular, we may make (1.5) hold.

It remains to show that the semisimple part of $\hat{T}$ is unique. Denote by $\mathfrak{F}_r$ the quotient algebra of $\mathfrak{F}$ modulo the ideal of all $\hat{f} \in \mathfrak{F}$ with $\pi_0 \hat{f} = \cdots = \pi_r \hat{f} = 0$. Denote by $\theta_r : \mathfrak{F} \to \mathfrak{F}_r$ the natural homomorphism, which may be regarded as a process of truncation of formal power series. Denote by $\mathfrak{G}_r$ the quotient group of $\mathfrak{G}$ over the normal subgroup of all $\hat{\sigma}$ with

$$\theta_r(\hat{\sigma}^* \hat{f}) = \theta_r \hat{f}, \qquad \hat{f} \in \mathfrak{F}.$$

Denote by $\theta_r' : \mathfrak{G} \to \mathfrak{G}_r$ the natural homomorphism. It is clear that, if $\theta_r' \hat{T} = \theta_r' \hat{U}$ for r arbitrarily large, then $\hat{T} = \hat{U}$. For any $\hat{f} \in \mathfrak{F}$, we define

$$(\theta_r' \hat{T})(\theta_r \hat{f}) = \theta_r(\hat{T} \hat{f}).$$

Through the above formula, each element of $\mathfrak{G}_r$ operates on $\mathfrak{F}_r$ as a nonsingular linear transformation. Since $J(\hat{T}) - J(\hat{\phi})$ is nilpotent, we conclude that $\theta_r' \hat{T} - \theta_r' \hat{\phi}$ is also nilpotent. On the other hand $\theta_r' \hat{T}$ and $\theta_r' \hat{\phi}$ commute. Therefore $\theta_r' \hat{\phi}$ is the uniquely determined semisimple part of $\theta_r' \hat{T}$ as a linear transformation of the finite dimensional space $\mathfrak{F}_r$. The above assertion is valid for any r. Hence $\hat{\phi}$ is unique.

For formal vector field $\hat{X} \in \mathfrak{X}$, the semisimple part $\hat{S}$ can be defined in the obvious way and has been shown to exist in [2]. Its uniqueness can also be verified in an analogous fashion as the above argument.

2. Construction of a local diffeomorphism with a given semisimple part.

THEOREM 2.1. *Let $\hat{\phi}$ be the semisimple part of $\hat{T} \in \mathfrak{G}$. Assume that $\hat{\phi}$ is linear and write $\phi = \hat{\phi}$ as an element of G. Then there exists $T' \in G$ such that ϕ is a semisimple part of T' and $\otimes T' = \hat{T}$.*

Proof. Let $\hat{\phi}$ be given as in Lemma 1.2.

Let $R_0{}^n$ be the vector space of the n-tuples of rational numbers over the field of the rational numbers R_0. Elements of $R_0{}^n$ will be written in the form $u = \{m_1(u), \cdots, m_n(u)\}$.

Let ν_i be one of the values of the logarithm of λ_i where λ_i, $i = 1, \cdots, n$, are the eigenvalues of $J(\hat{\phi})$. Let U be the subspace of $R_0{}^n$ such that $u \in U$ if and only if

$$(2.1) \qquad \begin{gathered} m_1(u) \operatorname{Re} \nu_1 + \cdots + m_n(u) \operatorname{Re} \nu_n = 0, \\ f(u) = (2\pi)^{-1} [m_1(u) \operatorname{Im} \nu_1 + \cdots + m_n(u) \operatorname{Im} \nu_n] \in R_0. \end{gathered}$$

We say that $u \in U$ is strongly integral if $f(u), m_1(u), \cdots, m_n(u)$,

$\cdots, m_n(u)$ are all integers. Denote by $U\langle\langle k\rangle\rangle$ the totality of strongly integral *elements of* U such that $m_k(u) \geqq -1$ and $m_i(u) \geqq 0$ for $i \neq k$.

Let u be an integral element of $R_0{}^n$ with $m_k(u) \geqq -1$ and $m_i(u) \geqq 0$ for $i \neq k$. Then $\lambda^u = 1$ if and only if $u \in U\langle\langle k\rangle\rangle$.

According to § 9, [2], there exist $u_1, \cdots, u_\lambda \in U^+$ and $v_1, \cdots, v_\mu \in U(k)$ such that any element u of $U\langle\langle k\rangle\rangle$ can be written in the form

$$u = \alpha_1 u_1 + \cdots + \alpha_\lambda u_\lambda + \beta_1 v_1 + \cdots + \beta_\mu v_\mu \tag{2.2}$$

for some non-negative rational numbers α's and β's with $\beta_1 + \cdots + \beta_\mu$ equal to either 0 or 1.

We may further assume that the u's are strongly integral. Define

$$\Phi = \sum_{i=1}^{\lambda} | z^{u_i} |^2.$$

Define $b_r{}^k(z)$ through

$$\pi_r \hat{T}^* z^k = b_r{}^k(z) z^k.$$

Then, since T and ϕ commute, we have, from Corollary, Lemma 1.3;

$$b_r{}^k(z) = \sum b_u{}^k z^u$$

where each $b_u{}^k$ is a complex number and the summation runs over all $u \in U\langle\langle k\rangle\rangle$ with $|u| = r-1$. Write

$$c_r = \sum | b_u{}^k |$$

summing over all $u \in U\langle\langle k\rangle\rangle$ with $|u| = r-1$ and over $k = 1, \cdots, n$. Set

$$h^k(z) = \sum_{r=0}^{\infty} b_r{}^k(z) z^k \chi(r : c_r \Phi).$$

We show that $h^k(z)$ is a complex function of C^∞ in R^n precisely in the same manner as in § 10, [2].

If z^i and z^j are conjugate, then $\pi_r T^* z^i$ and $\pi_r \hat{T}^* z^j$ must be conjugate. Consequently $h^i(z)$ and $h^j(z)$ must be conjugate. The C^∞ mapping T': $R^n \to C^n$ given through

$$T'^* z^i = h^i(z), \qquad i = 1, \cdots, n,$$

maps R^n into R^n. Moreover,

$$z^i \circ (T'\phi) = h^i(z \circ \phi) = \lambda^i h^i(z) = z^i \circ (\phi T'),$$

i.e. T' and ϕ commute. It is easy to show that $\Theta T' = \hat{T}$. Hence the theorem is proved.

Let J_+ (or J_-) be the subset of $\{1,\cdots,n\}$ that consists of all i with $|\lambda_i|>1$ (or <1). Denote by V^+ (or V^-) the subspace of R^n such that $z^i=0, i\in J_-$ (or $i\in J_+$). If $u\in U\langle\langle k\rangle\rangle$ and $k\in J_+$ (or $k\in J_-$) then $m_j(u)>0$ for some $j\in J_+$ (or $j\in J_-$), otherwise

$$m_j(u)\operatorname{Re}\nu_j=\sum_{j\in J_+} m_j(u)\operatorname{Re}\nu_j+\sum_{j\in J_-} m_j(u)\operatorname{Re}\nu_j$$

would not vanish. Therefore the function $b_r^k(z)z^k$ vanishes on V^- (or V^+), $r=0,1,\cdots$. We are led to the next assertion:

The C^∞ mapping T' constructed above leaves both V^+ and V^- invariant.

3. Theorems 1 and 2.

Proof of Theorems 1 *and* 2. Let $\hat{\phi}$ be the semisimple part of $\otimes T$. Assume that the coordinates x are so chosen that $\hat{\phi}$ is linear. Theorem 2.1 assures the existence of $T'\in G$ having the linear transformations $\hat{\phi}$ as its semisimple part and leaving V^+ and V^- invariant. Corollary, Lemma 3.1, [2], asserts that T and T' are equivalent. Thus T has a semisimple part ϕ. Evidently $\otimes\phi=\hat{\phi}$. The first half of Theorem 1 is proved.

The necessity part of Theorem 2 is trivial. For sufficiency, note that U is equivalent to some U' with $\otimes U'=\otimes T$. Now the construction of T' depends only on $\otimes T$, and U' and T' are also equivalent. Therefore, Theorem 2 is proved.

Finally, if ψ is also a semisimple part of T, then $\otimes\psi$ has to be $\hat{\phi}$. Clearly $\phi^{-1}\psi$ belongs to G_∞ and commutes with T. Conversely, let ψ be any local diffeomorphism such that this is true, then ψ commutes with T, and $J(\psi^{-1}T)-E=J(\phi^{-1}T)-E$ is nilpotent. Moreover, $\otimes\psi=\otimes\phi$ implies the equivalence of ψ and ϕ, and consequently ψ is also semisimple. Hence the proof is completed.

Example. Consider a local C^∞ diffeomorphism T about the origin of R^2 with the Jacobian matrix

$$J(T)=\operatorname{diag}(\lambda,\lambda^{-1}),$$

λ being real and $|\lambda|\neq 1$. By Theorem 1, there exists a local diffeomorphism σ such that the linear transformation ϕ given by $x\circ\phi=\lambda x$, $y\circ\phi=\lambda^{-1}y$, is a semisimple part of $U=\sigma T\sigma^{-1}$. Since ϕ leaves each curve of the family $xy=$ constant invariant, the commutativity of ϕ and U implies that U carries each of the hyperbolas to another.

4. Formal logarithm.

LEMMA 4.1. *Let $\hat{N}\in\mathfrak{A}$. If $\pi_0\hat{N}$, operating on $\mathfrak{F}^{(1)}$, is nilpotent, then,*

for any given positive integer r, there exists a sufficiently large positive integer s such that

$$\pi_r(\hat{N}^s\mathfrak{F}) = \{0\}.$$

Proof. Since, for $r = 0, 1, \cdots$, and $k > r$,

$$\pi_r(\hat{N}^s\mathfrak{F}^{(k)}) = \{0\}, \tag{4.1}$$

it suffices to show (4.1) for $k \leqq r$ and for s sufficiently large. Observe that $\pi_0\hat{N}$, operating on $\mathfrak{F}^{(k)}$, is nilpotent. It follows that, given k, there exists a positive integer s(k) such that

$$\pi_r(\hat{N}^{s(k)}\mathfrak{F}^{(k)}) = \{0\}.$$

We set $s = r + s(0) + \cdots + s(r)$. Thus the lemma is proved.

PROPOSITION 4.1. *If $\hat{\eta} \in \mathfrak{G}$ is such that $J(\hat{\eta})$ minus the identity linear transformation of R^n is nilpotent, then there exists a unique $\hat{N} \in \mathfrak{G}$ such that*

$$\hat{\eta} = \operatorname{Exp} \hat{N}$$

with $\pi_0\hat{N}$, operating on $\mathfrak{F}^{(1)}$, being nilpotent.

Proof. Write $\hat{N}_r = \pi_r\hat{N}$. Let E_r be the identity linear transformation of $\mathfrak{F}^{(r)}$. As a linear transformation of $\mathfrak{F}^{(1)}$, $\hat{N}_0$ must satisfy the condition

$$J(\hat{\eta}) = \operatorname{Exp} \hat{N}_0.$$

Since $\hat{N}_0$ is supposed to be nilpotent, we have

$$\hat{N}_0 = \log(E_1 + (J(\hat{\eta}) - E_1)) = (J(\hat{\eta}) - E_1) - (J(\hat{\eta}) - E_1)^2/2 + \cdots$$

which is uniquely determined.

We may assume that

$$\hat{N}_0 x^i = \epsilon_i x^{i-1}, \qquad i = 1, \cdots, n$$

where $\epsilon_i = 0$ or 1 with $\epsilon_1 = 0$. We are going to determine by induction,

$$\hat{N}_1 x^1, \cdots, \hat{N}_1 x^n, \cdots, \hat{N}_r x^1, \cdots, \hat{N}_r x^n, \cdots.$$

For $r \geqq 0$, we have

$$\begin{aligned}\pi_{r+1}\hat{\eta}^* x^i &= \pi_{r+1}(\exp \hat{N} x^i) \\ &= \pi_{r+1}(\hat{N}x^i + (2!)^{-1}\hat{N}^2 x^i + \cdots),\end{aligned}$$

which, according to Lemma 4.1, must be a finite sum and is equal to

$$\hat{N}_r x^i + (2!)^{-1} \sum_{j+k=r} \hat{N}_j \hat{N}_k x^i + (3!)^{-1} \sum_{j+k+l=r} \hat{N}_j \hat{N}_k \hat{N}_l x^i + \cdots.$$

Thus

$$\pi_{r+1}\hat{\eta}^* x^i = (E_{r+1} + (2!)^{-1}\hat{N}_0 + (3!)^{-1}\hat{N}_0^2 + \cdots)\hat{N}_r x^i$$
$$+ \text{ terms which can be determined by the induction hypothesis.}$$

Since the inverse of

$$E_{r+1} + (2!)^{-1}\hat{N}_0 + (3!)^{-1}\hat{N}_0^2 + \cdots$$

exists in $L(\mathfrak{F}^{(r+1)}, \mathfrak{F}^{(r+1)})$, $\hat{N}_r x^i$ is uniquely determined. Hence the proposition is proved.

PROPOSITION 4.2. *For $\hat{Y} \in \mathfrak{A}$, $\operatorname{Exp} \hat{Y} \in \mathfrak{G}$ is semisimple if and only if $\hat{Y}$ is semisimple.*

Proof. If $\hat{Y}$ is semisimple, then there exists $\hat{\sigma}$ such that $\hat{Y}' = \operatorname{Adj} \hat{\sigma}\hat{Y}$ is linear and diagonalizable. Then

$$\operatorname{Exp} \hat{Y}' = \operatorname{Exp}(\operatorname{Adj} \hat{\sigma}\hat{Y}) = \hat{\sigma}(\operatorname{Exp} \hat{Y})\hat{\sigma}^{-1}$$

is also linear and diagonalizable, i. e. $\operatorname{Exp} \hat{Y}$ is semisimple.

Conversely, let $\operatorname{Exp} \hat{Y}$ be semisimple and let $\hat{Y}'$ be the semisimple part of $\hat{Y}$. Then $\operatorname{Exp} \hat{Y}$ and $\operatorname{Exp} \hat{Y}'$ commute. Since $\pi_0\hat{Y}'$ is the semisimple part of $\pi_0\hat{Y}$, operating on $\mathfrak{F}^{(1)}$,

$$J(\operatorname{Exp} \hat{Y}') = \operatorname{Exp} \pi_0\hat{Y}'$$

is also the semisimple part of $J(\operatorname{Exp} \hat{Y})$. Therefore $\operatorname{Exp} \hat{Y}'$ is the semisimple part of $\operatorname{Exp} \hat{Y}$ in $\mathfrak{G}$ and is consequently identical with $\operatorname{Exp} \hat{Y}$. Write $N = \hat{Y} - \hat{Y}'$. Then

$$\operatorname{Exp} \hat{Y}' = \operatorname{Exp}(\hat{Y}' + N) = (\operatorname{Exp} \hat{Y})(\operatorname{Exp} N)$$

which implies $\operatorname{Exp} N = E$. It follows from Proposition 4.1 that $N = 0$. Hence $\hat{Y} = \hat{Y}'$ is semisimple.

THEOREM 4.1. *Let $\hat{\phi}$ be the semisimple part of $\hat{T}$. Then $\hat{T} = \operatorname{Exp} \hat{\mathfrak{X}}$ for some $\hat{\mathfrak{X}} \in \mathfrak{A}$ if and only if there exists $\hat{Y} \in \mathfrak{A}$ such that*

(a) $\hat{\phi} = \operatorname{Exp} \hat{Y}$,

(b) $\operatorname{Adj} \hat{T}\hat{Y} = \hat{Y}$, *i. e.* $\hat{T}^*\hat{Y} = \hat{Y}\hat{T}^*$.

Proof. Without loss of generality, we may assume $\hat{Y}$ to be linear.

Necessity: Let $\hat{Y}$ be the semisimple part of $\hat{\mathfrak{X}}$. Then the argument in

the necessity part of the preceding proposition may be used to show that $\operatorname{Exp}\hat{Y}$ is precisely the semisimple part $\hat{\phi}$ of $\hat{T}$. Since $\hat{X}$ and $\hat{Y}$ commute, we have

$$\pi_r\hat{T}^*\hat{Y}\hat{f} = \pi_r\hat{Y}\hat{f} + \cdots + \pi_r\hat{X}^r\hat{Y}\hat{f}/r! + \cdots = \pi_r\hat{Y}\hat{T}^*\hat{f}.$$

Therefore the condition is necessary.

Sufficiency: Write $\hat{\eta} = \hat{\phi}^{-1}\hat{T}$. Let $\hat{N}$ be given as in Proposition 4.1. Then

$$\hat{T} = (\operatorname{Exp}\hat{Y})(\operatorname{Exp}\hat{N}).$$

We set $\hat{X} = \hat{Y} + \hat{N}$. It remains to show that $[\hat{Y}, \hat{N}] = 0$. Observe that, owing to Remark, Lemma 1.1, we have

$$\begin{aligned}\hat{Y} &= \operatorname{Adj}\hat{T}\hat{Y} = (\operatorname{Adj}\hat{\eta})(\operatorname{Adj}\hat{\phi}\hat{Y}) = \operatorname{Adj}\hat{\eta}\hat{Y}\\ &= (\operatorname{Adj}\operatorname{Exp}\hat{N})Y = (\exp \operatorname{ad}\hat{N})\hat{Y}.\end{aligned}$$

Write

$$\hat{N}_r = \pi_r\hat{N}.$$

We are going to prove, by induction on r, that

$$[\hat{Y}, \hat{N}_r] = (\operatorname{ad}\hat{N}_r)\hat{Y} = 0, \qquad r \geqq 0.$$

Making use of the induction hypothesis when $r > 0$, we have for $r \geqq 0$

$$0 = \pi_r((\exp \operatorname{ad}\hat{N})\hat{Y} - \hat{Y}) = P(\operatorname{ad}\hat{N}_r)\hat{Y},$$

where

$$P = E_r' + (2!)^{-1}\operatorname{ad}\hat{N}_0 + (3!)^{-1}(\operatorname{ad}\hat{N}_0)^2 + \cdots$$

is a linear transformation of $\mathfrak{X}^{(r)}$ with E_r' being the identity transformation. Since $\operatorname{ad}\hat{N}_0$ is nilpotent, P is nonsingular. Hence $(\operatorname{ad}\hat{N}_r)\hat{Y} = 0$.

PROPOSITION 4.3. *Let $\{z^1, \cdots, z^n\}$ be a base of $\tilde{\mathfrak{F}}^{(1)}$ such that the conjugate of any z^i is also an element of the base. Let*

$$\hat{Y} = \sum \nu_i z^i \partial/\partial z^i$$

be real, i.e. $\hat{Y} \in \mathfrak{X}$. Then for any $\hat{T} \in \mathfrak{G}$, the following three conditions are equivalent:

(a) $\hat{T}^*\hat{Y}z^i = \hat{Y}\hat{T}^*z^i$, $i = 1, \cdots, n$.

(b) $\hat{T}^*\hat{Y} = \hat{Y}\hat{T}^*$.

(c) *For* $i = 1, \cdots, n$, *and* $r = 1, 2, \cdots$,

$$\pi_r\hat{T}^*z^i = \sum c_m{}^i z^m \tag{4.3}$$

where the c's are complex numbers, and the summation runs over all m with $|m| = r$ *and*

$$\sum m_j \nu_j = \nu_i. \tag{4.4}$$

Proof. If (a) holds, then, for any $\hat{f} \in \mathfrak{F}$,

$$\hat{Y}\hat{T}^*\hat{f} = \sum \nu_j z^j \partial \hat{T}^* \hat{f}/\partial z^j,$$

by chain rule of differentiation,

$$\begin{aligned} &= \sum_{i,j} \nu_j z^j (\hat{T}^* \partial \hat{f}/\partial z^i)(\partial \hat{T}^* z^i/\partial z^j) \\ &= \sum (\hat{Y}\hat{T}^* z^i)(\hat{T}^* \partial \hat{f}/\partial z^i) = \sum (\hat{T}^* \hat{Y} z^i)(\hat{T}^* \partial \hat{f}/\partial z^i) \\ &= \hat{T}_0 \hat{Y} \hat{f}. \end{aligned}$$

If (b) holds, let $\pi_r \hat{T}^* z^i = \sum b_m z^m$, summing over all m with $|m| = r$. Then

$$\nu_i \sum b_m z^m = \pi_r \hat{T}^* \hat{Y} z^i = \pi_r \hat{Y} \hat{T}^* z^i = \sum (\sum_j m_j \nu_j) b_m z^m.$$

Thus $b_m \neq 0$ implies (4.4), and (c) holds. Finally, let us assume (c). Then

$$\pi_r Y T^* z^i = \hat{Y} \sum c_m{}^i z^m = \nu_i \sum c_m{}^i z^m = \pi_r \hat{T}^* \hat{Y} z^i.$$

Hence the proof is completed.

THEOREM 4.2. *For any* $\hat{T} \in \mathfrak{G}$, *there exists a positive integer* α *such that*

$$\hat{T}^\alpha = \operatorname{Exp} \hat{X}$$

for some $\hat{X} \in \mathfrak{A}$.

Proof. Let $\hat{\phi}$ be the semisimple part $\hat{T}$, and assume that $\hat{\phi}$ is already in the form given in Lemma 1.2 such that

$$\hat{\phi}^* z^i = \lambda_i z^i.$$

We may further assume that none of the λ's is negative. Otherwise we may replace $\hat{T}$ by some power of $\hat{T}$. Since $\hat{T}$ and $\hat{\phi}$ commute, $\pi_r \hat{T}^* z^i$ is in the form given by Corollary, Lemma 1.3. It is clear that, for any positive integer l, $\hat{\phi}^l$ is the semisimple part of $\hat{T}^l$. Choose a value ν_i of $\log \lambda_i$ in such a way, if λ_i and λ_j are conjugate, so are ν_i and ν_j. We divide our discussion into two cases:

Case 1. Suppose that, for $i = 1, \cdots, n$, and for any n-tuple m of natural integers, $\lambda^m = \lambda_i$ implies (4.4). It follows from the preceding proposition that $\hat{T}^*$ and

$$\hat{Y} = \sum \nu_i z^i \partial/\partial z^i$$

commute. Then our theorem follows from Theorem 4.1.

Case 2. Suppose that, for some i and m, $\lambda^m = \lambda_i$ does not imply (4.4). Denote by V the vector space over R_0 generated by $\operatorname{Im} \nu_i/2\pi$, $i = 1, \cdots, n$. Note that $\lambda^m = \lambda_i$ implies $(\sum m_j \operatorname{Im} \nu_j - m_i \operatorname{Im} \nu_i)/2\pi$ being an integer. We claim that $R_0 \subset V$. For otherwise $\lambda^m = \lambda_i$ would always imply (4.4). There exists a subspace $V' \subset V$ such that

$$V = R_0 \oplus V'.$$

Let γ_i and δ_i be the projections of $\operatorname{Im} \nu_i/2\pi$ in R_0 and V' respectively such that $\operatorname{Im} \nu_i/2\pi = \gamma_i + \delta_i$, $i = 1, \cdots, n$. If ν_i and ν_j are conjugate, then $\gamma_i + \gamma_j = 0$. Let α be a positive integer such that each $\alpha\gamma_i$ is integral. We set

$$\nu_i' = \alpha(\nu_i - i2\pi\gamma_i) = \alpha(\operatorname{Re} \nu_i + i2\pi\delta_i)$$

and

$$\hat{Y} = \sum \nu_i z^i \partial/\partial z^i.$$

Then we have

$$\operatorname{Exp} \hat{Y} = \hat{\phi}^\alpha.$$

Since $\hat{T}^\alpha$ commutes with $\hat{\phi}$, Corollary, Lemma 1.3, holds for $\hat{T}^\alpha$. If $\lambda^m = \lambda_i$, then

$$\sum m_j \operatorname{Re} \nu_j' = \alpha \sum m_j \operatorname{Re} \nu_j = \operatorname{Re} \nu_i'.$$

Moreover,

$$\sum m_j \operatorname{Im} \nu_j' - m_i \operatorname{Im} \nu_i' = 2\pi\alpha(\sum m_j\delta_j - \delta_i)$$

must vanish. It follows from Proposition 4.4 that $\hat{T}^\alpha$ and $\hat{Y}$ commute. Hence the theorem is proved.

Proof of Theorem 3. The existence of $\log T$ obviously implies that of $\log \Theta T$. Conversely, if $\Theta T = \operatorname{Exp} \hat{X}$ for some $\hat{X} \in \mathfrak{A}$, then there exists $X \in A$ with $\Theta X = \hat{X}$, and we have

$$\Theta \operatorname{Exp} X = \operatorname{Exp} \hat{X} = \Theta T.$$

It follows from Theorem 2 that there exists $\sigma \in G$ with

$$T = \sigma(\operatorname{Exp} X)\sigma^{-1} = \operatorname{Exp}(\operatorname{Adj} \sigma X).$$

Hence we establish the first half of the theorem and, through Theorem 4.2, also the second half.

Let α be any positive integer. We are going to construct a local diffeomorphism T in R^3 such that $T, \cdots, T^{\alpha-1}$ do not lie on a 1-parameter group, but T^α does.

Example. Let R^3 be given the system of coordinates $(x, z, \bar{z})$, where x is real, and z and $\bar{z}$ are complex conjugate. Define the local diffeomorphism T such that

$$T^*x = ex,$$
$$T^*z = e^{-1+i2\pi\alpha^{-1}}z + x^\alpha z^{\alpha+1},$$
$$T^*\bar{z} = \text{the conjugate of } T^*z.$$

The semisimple part $\hat{\phi}$ of $\hat{T}$ has a logarithm

$$\hat{Y} = x\partial/\partial x + (-1 + i2\pi\alpha^{-1})z\partial/\partial z + (-1 - i2\pi\alpha^{-1})\bar{z}\partial/\partial\bar{z}$$

such that

$$\hat{Y}(x^\alpha z^\alpha) = i2\pi x^\alpha z^\alpha.$$

For any positive integer l, the semisimple part of $\hat{T}^l$ is $\hat{\phi}^l$, whose logarithms are

$$\hat{Y}_1 = l\hat{Y} + i2\pi(s_1 x\partial/\partial x + s_2 z\partial/\partial z + s_2\bar{z}\partial/\partial\bar{z})$$

where s_1 and s_2 are arbitrary integers. A straightforward computation shows that

$$T^{*l}x = e^l x,$$
$$T^{*l}z = (e^{-1+i2\pi\alpha^{-1}} + x^\alpha z^\alpha)^l z.$$

Now Y_1 and T^{*l} commute if and only if

$$\hat{Y}_1(x^\alpha z^\alpha) = 0$$

i.e.

$$i2\pi(l + \alpha(s_1 + s_2))x^\alpha z^\alpha = 0,$$

which may happen for suitable choices of s_1 and s_2 when and only when $l = 0$ mod α. For $0 \leqq l < \alpha$, this is possible only when $l = 0$. Hence α is the least positive integer such that $\log T^\alpha$ exists.

Remark 1. It can be easily seen through the theorems of this section that, for any diffeomorphism T about an elementary critical point in R^2, either $\log T$ or $\log T^2$ must exist.

Remark 2. In order to give a counter-example of a conjecture of Lie regarding the exponential map for an infinite Lie group, Sternberg shows in [10] to the effect that the local diffeomorphism U of R^2 (with the conjugate complex coordinates z and $\bar{z}$) given by

$$U^*z = e^{i2\pi\alpha^{-1}}z + az^{\alpha+1}$$

has no logarithm for some positive integer α and some nonzero complex number a with $|a|$ arbitrarily small. Here the fixed point is not elementary and therefore does not suit our purpose. However Sternberg's interesting example does suggest the kind of local diffeomorphisms to look for.

RUTGERS—THE STATE UNIVERSITY.

REFERENCES.

[1] K. T. Chen, "Decomposition of differential equations," *Mathematische Annalen*, vol. 146 (1962), pp. 263-278.

[2] ———, "Equivalence and decomposition of vector fields about an elementary critical point," *American Journal of Mathematics*, vol. 85 (1963), pp. 693-722.

[3] ———, "On local diffeomorphisms about an elementary fixed point," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 838-840.

[4] P. Hartman, "On local homeomorphisms of euclidean spaces," *Proceedings of the Symposium on Ordinary Differential Equations*, Mexico City, 1959.

[5] D. C. Lewis, "Invariant manifolds near an invariant point of unstable type," *American Journal of Mathematics*, vol. 60 (1938), pp. 577-587.

[6] ———, "On formal power series transformations," *Duke Mathematical Journal*, vol. 5 (1939), pp. 794-805.

[7] S. Smale, "Stable manifolds for differential equations and diffeomorphisms," *Annali della Scuola Normale Superiore de Pisa*, vol. 17 (1963), pp. 97-116.

[8] S. Sternberg, "Local contractions and a theorem of Poincaré," *American Journal of Mathematics*, vol. 79 (1957), pp. 809-824.

[9] ———, "The structure of local homeomorphism III," *American Journal of Mathematics*, vol. 81 (1959), pp. 578-604.

[10] ———, "Infinite Lie groups and the formal aspects of dynamic systems," *Journal of Mathematics and Mechanics*, vol. 10 (1961), pp. 451-474.

[11] P. Hartman, *Ordinary Differential Equations*, John Wiley, 1964.

HAUSDORFF MEASURE OF CRITICAL IMAGES ON BANACH MANIFOLDS.

By Arthur Sard.*

1. **Introduction.** Consider a map f of an m-dimensional manifold M into a Banach manifold N, where M is finite dimensional and has a countable basis, and f, M, N are of class C^q, $1 \leqq q \leqq \infty$. The manifolds M and N may include boundary; N need not have a countable basis.

In particular M may be an arbitrary open subset of Euclidean m-space E^m; N may be an arbitrary Banach space. A reader interested in this case may omit parts of our discussion.

The rank r of a point $x \in M$ is defined as the rank of the tangent map $Df(x)$; that is, r is the dimension of the image $Df(x)E^m$. Thus r is a non-negative integer $\leqq m$. If N is finite dimensional, then r equals the rank of the Jacobian matrix of f relative to local coordinates near x and $f(x)$. We put

$$k = m - r;$$

the integer k may be called the corank of x.

Let A_r denote the set of points of M of rank $\leqq r$. Then

$$A_0 \subset A_1 \subset \cdots \subset A_m = M.$$

We shall consider the image fA_r of A_r under f and particularly its Hausdorffian character, that is, whether it is s-null or s-sigmafinite for given s. In § 2 we discuss Hausdorff measure on Banach manifolds. Thereafter we establish the following theorem: if $\rho > 0$ and q is sufficiently large (depending on k and ρ), then fA_r is $(r+\rho)$-null.

It is essential here that the domain M of f be finite dimensional. Indeed Kupka [13] has constructed a C^∞-map $\bar{f}$ of a separable Hilbert space into E^1 which carries a set of points of rank 0 onto a set of positive 1-measure. Thus $\bar{f}A_r$ is not $(r+1)$-null, where $r = 0$.

The conclusion that fA_r is $(r+\rho)$-null for all maps $f \in C^q$ requires that $\rho > 0$. For, if $M = E^m$ and f is a projection in E^m onto an r-plane E^r, then $A_r = E^m$ and $fA_r = E^r$, which is r-positive and not r-null.

Received May 18, 1964.

* Research supported in part by the National Science Foundation.

Perhaps the most important cases are

$$0 < \rho \leqq 1$$

and particularly $\rho = 1$. If fA_r is $(r+\rho)$-null for some $\rho \leqq 1$, then a fortiori fA_r is $(r+1)$-null. Now the latter fact can be useful in many ways. For example it is readily seen that if M contains one point of rank $> r$ and if Ω is any neighborhood of that point, then $f\Omega$ is $(r+1)$-positive (Cf. Lemma 5 below.). Thus the existence of one point of rank $> r$ implies that the entire image fA_r of points of rank $\leqq r$ is negligible.

The absence of a Lebesgue measure in infinite dimensional space is no barrier to our analysis. Hausdorff measure is available and pertinent.

Known facts and open questions related to critical images are discussed in § 5. My abstract [25] referred to an earlier, unpublished version of the present paper, with different title.

2. Hausdorff measure on Banach manifolds. Consider first a Banach space Y. Suppose that B is a subset of Y. We shall define the s-dimensional measure $\mu_s B$ of B.

The 0-dimensional measure of B is simply the number of points in B.

Let $s > 0$. We say that B is *s-null* if for each $\epsilon > 0$ there is a countable cover of B by sets for which the sum of the s-th powers of the diameters of the covering sets is less than ϵ. To define $\mu_s B$ for any $B \subset Y$, we first define the coarse measure $\mu_{s,\zeta} B$ as follows. For each $\zeta > 0$, $\mu_{s,\zeta} B$ is the infimum of the sum of the s-th powers of the diameters of the covering sets, among countable covers of B by sets of diameter $< \zeta$. If there is no such cover, $\mu_{s,\zeta} B = \infty$. Then $\mu_{s,\zeta} B$ is a monotone function of ζ, and $\mu_s B$ is defined as the limit (possibly infinite) of $\mu_{s,\zeta} B$ as $\zeta \to 0+$. One may show that $\mu_s B = 0$ if and only if B is s-null as earlier defined. It is sometimes convenient to multiply the powers of the diameters by a normalizing factor depending on s [23]; we will not do so. The above definitions apply more generally to the case in which Y is a metric space.

We say that B is *s-positive*, *s-finite*, or *s-infinite* according as $\mu_s B$ has each property. We say that B is *s-sigmafinite* if B is the countable union of s-finite sets.

Any subset of an s-null set is s-null. Any countable union of s-null sets is s-null. Likewise any subset of an s-sigmafinite set is s-sigmafinite. Any countable union of s-sigmafinite sets is s-sigmafinite.

If $B \subset Y$ is s-sigmafinite, then B is $(s+\rho)$-null for all $\rho > 0$.

LEMMA 1. *Suppose that θ is a differentiable map of an open subset Ω of*

a Banach space into a Banach space. If $A \subset \Omega$ is s-null or s-sigmafinite, then so is θA.

Proof. For each pair ν, λ of positive integers, let $\Omega^{\nu,\lambda}$ be the set of points $x \in \Omega$ such that

$$(1) \quad \| \theta(x') - \theta(x) \| < \nu \| x' - x \| \text{ whenever } \| x' - x \| < 1/\lambda, \quad x' \in \Omega.$$

There are countably many sets $\Omega^{\nu,\lambda}$; and

$$\Omega = \bigcup_{\nu,\lambda} \Omega^{\nu,\lambda},$$

because the differentiability of θ at $x \in \Omega$ implies that (1) must hold for some ν and some λ.

If A is s-null, so is its subset $A \cap \Omega^{\nu,\lambda}$ and therefore $\theta(A \cap \Omega^{\nu,\lambda})$ by (1) and therefore $A = \bigcup_{\nu,\lambda} \theta(A \cap \Omega^{\nu,\lambda})$.

Similarly if A is s-sigmafinite. This completes the proof.

Observe that we have not restricted the Banach spaces that carry the domain and range of θ in any way.

Lemma 1 implies that the Hausdorffian character of sets in a Banach space is unchanged if the norm is replaced by an equivalent norm, as is evident directly.

Now consider a C^q-Banach manifold N, $q \geqq 1$. Detailed definitions are given in [18, § 3], [14, Ch. 2]; cf. [17] and [20] also. The manifold N is a topological space together with a differential structure. A tool for studying N is its complete atlas, which is a set of charts. Each chart (V, ψ, Y) is an open set $V \subset N$, a Banach space or half Banach space Y, and a homeomorphism ψ of V onto the open set $\psi V \subset Y$. If (V, ψ, Y) and (V', ψ', Y') are charts for N, then $\psi'\psi^{-1}$ is a C^q-diffeomorphism of $\psi(V \cap V') \subset Y$ onto $\psi'(V \cap V') \subset Y'$. It is understood that if $\psi(V \cap V')$ is not open in a full Banach space, then $\psi'\psi^{-1}$ can be extended so as to be together with its inverse C^q on an open set of a full Banach space. One defines what is meant by a C^q-map and a C^q-diffeomorphism of an open subset of N. In particular if (V, ψ, Y) is a chart for (part of) N, then ψ is a C^q-diffeomorphism of V. We sometimes describe the chart (V, ψ, Y) briefly as ψ.

For the present paper N need not have a countable basis, nor be a Hausdorff space, nor be paracompact.

Suppose that $s > 0$ and that $B \subset N$. We say that B is ***s*-null** if there exist countably many charts (V^i, ψ^i, Y^i) of N such that $B \subset \bigcup_i V^i$ and $\psi^i(B \cap V^i)$ is s-null in the space Y^i for all $i = 1, 2, \cdots$. Similarly we say

that B is *s-sigmafinite* if $\psi^i(B \cap V^i)$ is s-sigmafinite for all $i=1,2,\cdots$. We say that B is *s-positive* if B is not s-null.

LEMMA 2. *Any subset of an s-null set on N is s-null. Likewise a subset of an s-sigmafinite set is s-sigmafinite.*

Proof. Immediate.

LEMMA 3. *A countable union of s-null sets on N is s-null. Likewise a countable union of s-sigmafinite sets is s-sigmafinite.*

Proof. Suppose that

$$B = \bigcup_{\nu=1,2,\cdots} B^\nu \subset N \tag{2}$$

and that charts $(V^{\nu,i}, \psi^{\nu,i}, Y^{\nu,i})$ exist such that $B^\nu \subset \bigcup_{i=1,2,\cdots} V^{\nu,i}$ and $\psi^{\nu,i}(B^\nu \cap V^{\nu,i})$ is s-null in $Y^{\nu,i}$; ν, $i=1,2,\cdots$. Then $B \subset \bigcup_{\nu,i} V^{\nu,i}$ and therefore B is surely s-null if $\psi^{\nu,i}(B \cap V^{\nu,i})$ is s-null, all ν, i. For the latter it is sufficient that $\psi^{\nu,i}(B^\lambda \cap V^{\nu,i})$ is s-null, $\lambda=1,2,\cdots$, by (2). For the latter, in turn, it is sufficient that $\psi^{\nu,i}(B^\lambda \cap V^{\nu,i} \cap V^{\lambda,j})$ is s-null, $j=1,2,\cdots$, since

$$B^\lambda \cap V^{\nu,i} \subset \bigcup_{j=1,2,\cdots} B^\lambda \cap V^{\nu,i} \cap V^{\lambda,j}.$$

We therefore consider $\psi^{\nu,i}(B^\lambda \cap V^{\nu,i} \cap V^{\lambda,j})$. Now $\psi^{\lambda,j}(\psi^{\nu,i})^{-1}$ is a diffeomorphism on $\psi^{\nu,i}(V^{\nu,i} \cap V^{\lambda,j})$ onto $\psi^{\lambda,j}(V^{\nu,i} \cap V^{\lambda,j})$. And $\psi^{\lambda,j}(B^\lambda \cap V^{\lambda,j})$ is s-null by construction. Hence $\psi^{\lambda,j}(B^\lambda \cap V^{\nu,i} \cap V^{\lambda,j})$ is s-null, by Lemma 2. Hence $\psi^{\nu,i}(B^\lambda \cap V^{\nu,i} \cap V^{\lambda,j})$ is s-null by Lemma 1.

Thus B is indeed s-null.

A similar argument applies if B is a countable union of s-sigmafinite sets.

LEMMA 4. *If $B \subset N$ is s-sigmafinite, then B is $(s+\rho)$-null for all $\rho > 0$.*

Proof. Immediate, from the corresponding property for subsets of a Banach space.

Lemma 1 implies the following.

LEMMA 1′. *Suppose that θ is a C^q-map, $q \geqq 1$, of an open subset Ω of a Banach manifold into a Banach manifold. If $A \subset \Omega$ is s-null or s-sigmafinite, then so is θA.*

LEMMA 5. *If $B \subset N$ and if one chart (V, ψ, Y) for N exists such that $\psi(B \cap V)$ is s-positive in Y, then B is s-positive in N.*

Proof. Otherwise B would be s-null in N, and $B \cap V$ also, by Lemma 2. Then $\psi(B \cap V)$ would be s-null in Y, by Lemma 1′. This contradiction establishes the lemma.

It may be of interest to describe the Hausdorff dimension of $B \subset N$.

It is the infimum of numbers s for which B is s-null. Thus each of the following statements implies its successor:

1. B is s-null.
2. B is s-sigmafinite.
3. B is of Hausdorff dimension $\leqq s$, that is, B is $(s+\rho)$-null for all $\rho > 0$.

Besicovitch [1] has shown that 3) does not imply 2).

3. Images of critical sets of differentiable maps. Define the function c as follows:

$$(3) \qquad c(k,\rho) = \begin{cases} 1 & \text{if } k < \rho, \\ 2 + k(k-\rho)/\rho & \text{if } 0 < \rho \leqq k, \end{cases} \quad k = 0,1,2,\cdots; \quad \rho > 0.$$

Suppose that f is a C^q-map of a C^q-manifold M of dimension $m < \infty$ with countable basis into a C^q-Banach manifold N, where $1 \leqq q \leqq \infty$. Let A_r denote the set of points of M of rank $\leqq r$, $r = 1,2,\cdots,m$. Put

$$k = m - r.$$

THEOREM 1. *Suppose that* $\rho > 0$. *Then* fA_r *is* $(r+\rho)$*-null if*

$$q \geqq c(k,\rho).$$

Proof. The manifold M, being of dimension m with countable basis, is m-sigmafinite. Hence the entire image fM is m-sigmafinite, by Lemma 1′. It follows that fM and fA_r are $(r+\rho)$-null if $r+\rho > m$, by Lemmas 4 and 2.

Let $\nu(k,\rho)$ denote the smallest integer $> k/\rho$:

$$(4) \qquad \nu(k,\rho) - 1 \leqq k/\rho < \nu(k,\rho).$$

The properties of the function $c(k,\rho)$ which will enter in our proof are the following: $c(k,\rho)$ is defined for $k = 0,1,2,\cdots$ and for $\rho > 0$; for all such k, ρ,

$$(5) \qquad c(k+1,\rho) \geqq c(k,\rho) \geqq 1,$$

$$(6) \qquad c(k,\rho) \geqq \nu(k,\rho),$$

$$(7) \qquad c(k,\rho) \geqq c(k-1,\rho) + \nu(k,\rho) - 2 \text{ whenever } 0 < \rho \leqq k \text{ and } \nu(k,\rho) \geqq 3.$$

$$(8) \qquad c(k,\rho) \geqq c(k+1,\rho+1) \text{ whenever } 0 < \rho \leqq k.$$

The reader may verify that the function (3) satisfies these conditions. Indeed it satisfies the conditions with $\nu(k,\rho)$ replaced by $1 + k/\rho$.

If $x \in M$, then the rank of x is well defined, whether x is an interior point or boundary point of M [17, p. 5], [14, p. 31], [18, § 4]. Let B_r denote the set of points of M of rank r. Thus

$$A_r = B_0 \cup B_1 \cup \cdots \cup B_r, \qquad r = 0, 1, \cdots, m.$$

If the manifold N is of dimension $n < m$, then $B_{n+1}, \cdots, B_m$ are empty and the points of B_n are the regular points of the map f. If N is of dimension $\geqq m$, then the points of B_m are the regular points. If r is less than m and also less than $\dim N$, then A_r consists entirely of critical points; the lower r, the more highly critical, in a manner of speaking.

To prove that fA_r is $(r+\rho)$-null it will be sufficient to prove that fB_r is $(r+\rho)$-null whenever $q \geqq c(k, \rho)$. For, by the latter result with r replaced by $r-1$ and ρ replaced by $\rho+1$, because of (8), it follows that $q \geqq c(k+1, \rho+1)$ and that fB_{r-1} is $(r-1+\rho+1)$-null, that is, $(r+\rho)$-null. Similarly $fB_{r-2}, fB_{r-3}, \cdots, fB_0$ are all $(r+\rho)$-null, and therefore

$$fA_r = fB_0 \cup fB_1 \cup \cdots \cup fB_r$$

is $(r+\rho)$-null.

To prove that fB_r is $(r+\rho)$-null it will be sufficient to show that each point of B_r lies in an open set Ω such that $f(B_r \cap \Omega)$ is $(r+\rho)$-null. For, countably many such open sets will then cover B_r, since M has a countable basis; and fB_r will therefore be $(r+\rho)$-null by Lemma 3.

It will be sufficient to suppose that the manifold M is homogeneous, that is, everywhere of dimension m. For M may be divided into $m+1$ homogeneous pieces of dimension $m, m-1, \cdots, 2, 1, 0$, respectively, to each of which the theorem for homogeneous source may be applied, because of (5).

The remainder of the proof consists of four parts.

Part 1. *Local analysis of* f. Consider a point $u_0 \in B_r$. Suppose first that u_0 is not a boundary point of M and that $f(u_0)$ is not a boundary point of N. Choose charts (U, ϕ, X) for M and (V, ψ, Y) for N such that

$$u_0 \in U, \ f(u_0) \in V, \ \phi(u_0) = 0 \in X = E^m, \ \psi f(u_0) = 0 \in Y.$$

Here U, V are open in M, N, respectively; ϕ and ψ are C^q-diffeomorphisms; and ϕU, ψV are open in X, Y, respectively. Put

$$g = \psi f \phi^{-1}. \tag{9}$$

Then g is a C^q-map of ϕU into ψV. Note that $q \geqq 1$ by (5).

Let

$$t = Dg(0)$$

be the map tangent to g at 0. Put

$$Y_1 = tX \subset Y;$$

t is a linear continuous map of X onto Y_1. Since u_0 is of rank r, Y_1 is a linear space of dimension $r \leqq m < \infty$.

It follows that there exists a topological complement Y_2 of Y_1 in Y [9, pp. 347, 480], [4, pp. 98-99]. Then each $y \in Y$ may be written

$$y = y_1 + y_2, \qquad y_1 \in Y_1, \qquad y_2 \in Y_2,$$

in unique fashion; and the operators Q_1 and Q_2 defined by the relations

$$Q_1 y = y_1, \qquad Q_2 y = y_2, \qquad y \in Y,$$

are linear and continuous on Y.

Let X_2 be the kernel of t, that is,

$$X_2 = \{x \in X : t(x) = 0\};$$

and put

$$X_1 = X_2{}^{\perp}.$$

Let $\tau = t : X_1$ be the restriction of t to X_1. Then

$$\tau X_1 = Y_1 = tX, \qquad X_1 = \tau^{-1} Y_1, \qquad tX_2 = 0;$$

X_1 and Y_1 are linearly homeomorphic; and

$$\dim X_1 = \dim Y_1 = r, \qquad \dim X_2 = m - r = k.$$

We now introduce a new chart for M near u_0. To this end consider the map σ of $\phi U \subset X$ into X defined as follows:

$$\sigma(w) = (x_1, x_2), \quad w \in \phi U \subset X,$$

where

$$\begin{aligned} x_1 &= \tau^{-1} Q_1 g(w) \in X_1, \\ x_2 &= \operatorname{Proj}_{X_2} w \in X_2. \end{aligned} \tag{10}$$

Note that σ is well defined, since $Q_1 g(w) \in Y_1 =$ domain of τ^{-1}. The map σ is of class C^q since τ^{-1}, Q_1, $\operatorname{Proj}_{X_2}$ are linear continuous operators.

The map $D\sigma(0)$ tangent to σ at 0 carries $w' \in X$ into (x_1', x_2'), where

$$\begin{aligned} x_1' &= \tau^{-1} Q_1 Dg(0)(w') = \tau^{-1} Q_1 t(w') \\ &= \tau^{-1} Q_1 t(\operatorname{Proj}_{X_1} w' + \operatorname{Proj}_{X_2} w') = \tau^{-1} Q_1 \tau \operatorname{Proj}_{X_1} w' \\ &= \tau^{-1} \tau \operatorname{Proj}_{X_1} w' = \operatorname{Proj}_{X_1} w', \\ x_2' &= \operatorname{Proj}_{X_2} w', \end{aligned}$$

since t vanishes on X_2 and equals τ on X_1 and since $Q_1 \tau = \tau$. Thus $D\sigma(0)$ is the identity on X.

We may therefore apply the inverse function theorem to σ near 0. There is an open neighborhood $W \subset \phi U \subset X$ of 0 on which σ is a C^q-diffeomorphism onto $\sigma W \subset X$, and $\sigma(0) = 0$. Let

$$I = I_1 \times I_2$$

be an open interval centered at 0, contained with its closure in σW, and such that I_1 is a cubical interval (has equal edges) in X_1 and I_2 is a cubical interval in X_2. Then

$$(\phi^{-1}\sigma^{-1}I, \sigma\phi, X)$$

is an admissible chart for M near u_0. That is, $\sigma\phi$ is a C^q-diffeomorphism of the open set $\phi^{-1}\sigma^{-1}I \subset M$ onto the open set $I \subset X$ and $u_0 \in \phi^{-1}\sigma^{-1}I$.

Consider the C^q-map

$$h = g\sigma^{-1} = \psi f \phi^{-1}\sigma^{-1} = \psi f(\sigma\phi)^{-1}$$

on I into ψV. The map h is the local representation of f in terms of the charts $\sigma\phi$ of M and ψ of N. An arbitrary point $u \in \phi^{-1}\sigma^{-1}I$ is of rank r relative to f if and only if $\sigma\phi(u) \in I$ is of rank r relative to h. Let S denote the set of points of I of rank r relative to h. Then $\phi^{-1}\sigma^{-1}S$ is precisely the set of points of $\phi^{-1}\sigma^{-1}I$ of rank r relative to f; that is,

$$\phi^{-1}\sigma^{-1}S = B_r \cap \phi^{-1}\sigma^{-1}I.$$

We have seen that it will be sufficient for our theorem to show that there is a neighborhood Ω of each $u_0 \in B_r$ such that $f(B_r \cap \Omega)$ is $(r+\rho)$-null. Taking

$$\Omega = \phi^{-1}\sigma^{-1}I,$$

we see that it will therefore be sufficient to show that hS is $(r+\rho)$-null, by Lemma 1′, since

$$hS = h\sigma\phi(B_r \cap \Omega) = \psi f(B_r \cap \Omega)$$

and since ψ is a diffeomorphism.

The map h carries (x_1, x_2), $x_1 \in I_1$, $x_2 \in I_2$, into (y_1, y_2), where

$$\begin{aligned} y_1 &= Q_1 h(x_1, x_2) \in Y_1, \\ y_2 &= Q_2 h(x_1, x_2) \in Y_2, \end{aligned}$$

and Q_1, Q_2 are the projections defined earlier. Now by (10),

$$\begin{aligned} Q_1 h(x_1, x_2) = Q_1 g\sigma^{-1}(x_1, x_2) = Q_1 g(w) &= \tau(x_1), \\ (x_1, x_2) &= \sigma(w), w \in \sigma^{-1}I \subset W, \end{aligned}$$

since σ is a diffeomorphism on W. Thus h is the map

$$(11) \qquad \begin{aligned} y_1 &= \tau(x_1), \\ y_2 &= j(x_1, x_2), \end{aligned} \qquad x_1 \in I_1, x_2 \in I_2;$$

where

$$j(x_1, x_2) = Q_2 h(x_1, x_2) \in Y_2.$$

The map h is partially linear. For all (x_1, x_2) with $x_1 \in I_1$, $x_2 \in I_2$, the

component y_1 is a toplinear image of x_1 and is independent of x_2. The map j is C^q on the closure of I into Y_2.

For each $x_1 \in I_1$ and $x_2 \in I_2$, the derivative $Dh(x_1, x_2)$ is the map

$$(12) \qquad \begin{aligned} y_1' &= \tau(x_1'), \\ y_2' &= D_1 j(x_1, x_2)(x_1') + D_2 j(x_1, x_2)(x_2'), \end{aligned} \qquad x_1' \in X_1, x_2' \in X_2,$$

of X into Y, where D_1 indicates the partial derivative with respect to the first argument, $D_1 j(x_1, x_2)$ is a linear map of X_1 into Y_2, and D_2 is similar. Since Y_1 is of dimension r, it follows that for any $(x_1, x_2) \in I$, the image $Dh(x_1, x_2)X$ is r-dimensional if and only if

$$D_2 j(x_1, x_2) = 0.$$

Hence

$$(13) \qquad S = \{(x_1, x_2) \in I : D_2 j(x_1, x_2) = 0\};$$

that is, $(x_1, x_2) \in I$ is of rank r relative to h if and only if x_2 is of rank 0 relative to the map j with x_1 fixed; or equivalently, x_2 is of rank 0 relative to the map h with x_1 fixed.

The above analysis, in a simpler form, applies if $r = m$. In this case $h = \tau$ and all points of I are of rank m.

We have now completed the local analysis of f near u_0 in the case in which u_0 and $f(u_0)$ are interior points of M, N, respectively. If u_0 or $f(u_0)$ were a boundary point we could, by the definition of a C^q-map, extend the local description h of f to a map $\tilde{h}$, say, which is C^q on an open interval of E^m into an entire Banach space. Since the image under $\tilde{h}$ contains that under h, it will be sufficient to consider the former. This we shall do. For simplicity we designate $\tilde{h}$ simply as h. Thus the form (11) applies in all cases.

We proceed now to prove that hS is $(r + \rho)$-null.

Part 2. Separation of cases. Suppose that $r = \dim Y < \infty$, Then Y, being an r-dimensional Banach space, is r-sigmafinite, by a theorem of Tynchonoff. Hence Y and a fortiori hI and hS are $(r + \rho)$-null, for all $\rho > 0$, by Lemmas 4 and 2.

As we have already settled the cases $r + \rho > m$ at the beginning of the proof, we now assume that

$$(14) \qquad r < m, \qquad r < \dim Y, \qquad 0 < \rho \leqq m - r = k.$$

Divide the set S into subsets Z and T as follows:

$$\begin{aligned} Z = \{(x_1, x_2) \in S : D_2 j(x_1, x_2) = 0 \text{ and } D_2{}^2 j(x_1, x_2) = 0 \\ \text{and } \cdots \text{ and } D_2{}^{\nu-1} j(x_1, x_2) = 0\}, \end{aligned}$$

$$T = S - Z,$$

where $\nu = \nu(k, \rho)$ and the integer $\nu(k, \rho)$ is determined by (4). Since $\rho \leqq k$ by (14), it follows that $\nu \geqq 2$. If $\nu = 2$, then $Z = S$ by (13) and $T = 0$. If $\nu > 2$, then Z consists of those points of S at which not only the first partial of j relative to $x_2 \in X_2$ but all partials of j, relative to $x_2 \in X_2$ alone, of order $< \nu$ vanish.

Then

$$hS = hZ \cup hT.$$

To complete the proof we show that hZ is $(r+\rho)$-null (by a covering argument) and hT is $(r+\rho)$-null (by an induction.)

Part 3. *hZ is $(r+\rho)$-null.* Taylor's formula applied to the map j, the definition of Z, and (6) imply that there are constants α_1 and β_1 such that

$$\| j(x_1', x_2') - j(x_1, x_2) \|_{Y_2} \leqq \alpha_1 \| x_1' - x_1 \|_{X_1} + \beta_1 \| x_2' - x_2 \|_{X_2}{}^{\nu}$$
$$\text{whenever } (x_1, x_2) \in Z \text{ and } (x_1', x_2') \in I,$$

since

$$j(x_1', x_2') - j(x_1, x_2) = j(x_1', x_2') - j(x_1, x_2') + j(x_1, x_2') - j(x_1, x_2).$$

It follows from (11), since τ is a bounded operator, that for suitable constant α_2,

$$\| h(x_1', x_2') - h(x_1, x_2) \|_{Y} \leqq \alpha_2 \| x_1' - x_1 \|_{X_1} + \beta_1 \| x_2' - x_2 \|_{X_2}{}^{\nu} \tag{15}$$
$$\text{whenever } (x_1, x_2) \in Z \text{ and } (x_1', x_2') \in I.$$

We shall divide $I = I_1 \times I_2$ into small subintervals of such shape as to take advantage of the ratio $1/\nu$ of exponents in (15). Let a be the length of the edge of the cube I_1 and b that of I_2. Divide each edge of I_1 into 2^{λ} equal parts and each edge of I_2 into 2^{γ} equal parts, where λ, γ are positive integers to be specified later. Then I is divided into

$$2^{\lambda r + \gamma k}$$

congruent subintervals, in each of which each edge parallel to an edge of I_1 is of length $a/2^{\lambda}$, of I_2 of length $b/2^{\gamma}$. By (15) the image under h of any subinterval that intersects Z has diameter δ such that

$$\delta \leqq \frac{\alpha}{2^{\lambda}} + \frac{\beta}{2^{\gamma\nu}},$$

where α, β are suitable constants independent of λ, γ.

Thus the images of subintervals which intersect Z constitute a cover of hZ for which

$$\sum \delta^{r+\rho} \leqq 2^{\lambda r + \gamma k} \left(\frac{\alpha}{2^{\lambda}} + \frac{\beta}{2^{\gamma\nu}} \right)^{r+\rho} = \left(\frac{\alpha}{2^{e}} + \frac{\beta}{2^{e'}} \right)^{r+\rho},$$

where

$$e = \lambda - \frac{\lambda r + \gamma k}{r + \rho} = \lambda \frac{\rho}{r+\rho} - \gamma \frac{k}{r+\rho},$$

$$e' = \gamma \nu - \frac{\lambda r + \gamma k}{r+\rho} = -\lambda \frac{r}{r+\rho} + \gamma\left(\nu - \frac{k}{r+\rho}\right).$$

Now

$$\nu > \frac{k}{\rho}$$

by (4). It follows that we may choose a sequence of values of (λ, γ) such that

$$(\lambda, \gamma) \to (\infty, \infty), \quad e \to \infty, \quad e' \to \infty.$$

This is because there is a sector of positive angle in the first quadrant of the λ, γ-plane in which e and e' are always positive. Since α, β, and $r + \rho$ are constant, it follows that hZ is $(r+\rho)$-null.

Part 4. hT is $(r+\rho)$-null. We have seen in Part 2 that $T = 0$ if $\nu = 2$. It will be sufficient therefore to consider hT in the cases

$$\nu \geqq 3, \qquad 0 < \rho \leqq k,$$

by (14). In particular $k \geqq 1$, since k is an integer.

The set T consists of points $(x_1, x_2) \in I$ at which $D_2 h(x_1, x_2) = 0$ but not all maps $D_2{}^2 h(x_1, x_2)$, $D_2{}^3 h(x_1, x_2), \cdots, D_2{}^{\nu-1} h(x_1, x_2)$ vanish. Divide T into $\nu - 2$ subsets as follows:

$$T_i = \{(x_1, x_2) \in S\colon D_2 h(x_1, x_2) = 0 \text{ and } \cdots$$
$$\text{and } D_2{}^i h(x_1, x_2) = 0 \text{ and } D_2{}^{i+1} h(x_1, x_2) \neq 0\}, \quad i = 1, 2, \cdots, \nu - 2.$$

Then

$$hT = hT_1 \cup hT_2 \cup \cdots \cup hT_{\nu-2},$$

and it will be sufficient to prove that each term on the right is $(r+\rho)$-null.

Consider hT_i, with i fixed and $1 \leqq i \leqq \nu - 2$. To prove that hT_i is $(r+\rho)$-null it will be sufficient to show that each point $x_0 \in T_i$ lies in an open set Ω such that $h(T_i \cap \Omega)$ is $(r+\rho)$-null, since countably many such open sets will cover T_i. By the lemma on reduction of dimension which we state and prove in the next section, with $\theta = D_2{}^i h$, there is an open neighborhood Ω of $x_0 \in T_i$ in which one coordinate of x, say x^m, satisfies a relation

$$x^m = \eta(x^1, \cdots, x^{m-1}) \text{ whenever } x = (x^1, \cdots, x^m) \in T_i \cap \Omega,$$

where η is a $C^{q-\nu+2}$-map of an open neighborhood $\bar{\Omega}$ of $(x_0{}^1, \cdots, x_0{}^{m-1}) \in E^{m-1}$. The stated class of η is assured because $D_2{}^i h$ is a C^{q-i}-map of I into a Banach space of maps and because $i \leqq \nu - 2$.

It follows that $h(T_i \cap \Omega)$ is contained in $\tilde{h}\tilde{\Omega}$, where

$$\tilde{h}(x^1, \cdots, x^{m-1}) = h[x^1, \cdots, x^{m-1}, \eta(x^1, \cdots, x^{m-1})], \; (x^1, \cdots, x^{m-1}) \in \tilde{\Omega}.$$

More than this: $h(T_i \cap \Omega)$ is contained in the image under $\tilde{h}$ of the points of $\tilde{\Omega}$ of rank $\leqq r$ relative to $\tilde{h}$, because the rank of $(x^1, \cdots, x^{m-1})$ relative to $\tilde{h}$ is $\leqq$ the rank of $[x^1, \cdots, x^{m-1}, \eta(x^1, \cdots, x^{m-1})]$ relative to h. Thus $h(T_i \cap \Omega)$ is surely $(r+\rho)$-null if $\tilde{h}\tilde{S}$ is $(r+\rho)$-null, where $\tilde{S}$ consists of the points of $\tilde{h}$ of rank $\leqq r$.

Consider the case $m = 1$. Here the lemma on reduction of dimension implies that x_0 is isolated in $T_i \cap \Omega$. Thus $h(T_i \cap \Omega)$ consists of a single point and is therefore $(r+\rho)$-null. Thus Part 4 is completed and our entire theorem is proved in the case $m = 1$. The theorem is evident in the case $m = 0$, since M is then a countable set.

Now assume that the theorem has been established for maps on manifolds of dimension $m-1$ into N. Then $\tilde{h}\tilde{S}$ is surely $(r+\rho)$-null, since $\tilde{h}$ is such a map and since

$$q - \nu + 2 \geqq c(k-1, \rho),$$

by (7). Hence $h(T_i \cap \Omega)$ is $(r+\rho)$-null, and the theorem is established for maps on manifolds of dimension m.

This completes the proof. Similar inductions have been used by A. B. Brown [2], M. Morse and Sard [16], and Pontryagin [19].

COROLLARY. *If f is a C^∞-map, then fA_r is $(r+\rho)$-null for all $\rho > 0$.*

This corollary, at least in the case in which N is finite dimensional, is due to Dubovickii [8].

Remark. It may be of interest to note the following elementary theorem. Let us say that a point $u \in M$ is *simple* if f is of constant rank in some neighborhood of u. Let K denote the set of simple points of M. Then *the image $f(A_r \cap K)$ of simple points of rank $\leqq r$ is r-sigmafinite.* No condition on q other than $q \geqq 1$ is needed. To prove the theorem it is sufficient to show that $f(B_r \cap K)$ is r-sigmafinite. If $u_0 \in B_r \cap K$, then the neighborhood $I_1 \times I_2$ of Part 1 above may be taken so that $D_2 j(x_1, x_2)$ depends on x_1 alone, so that

$$j(x_1, x_2) = j(x_1), \qquad x_1 \in I_1, x_2 \in I_2.$$

Introduce the new chart $\omega\psi$ for N, where ω is the diffeomorphism

$$(z_1, z_2) = \omega(y_1, y_2),$$
$$z_1 = y_1,$$
$$z_2 = y_2 - j\tau^{-1}(y_1), \qquad y_1 \in \tau I_1, y_2 \in Y_2.$$

Then the local representation of the map f in terms of the charts $\sigma\phi$ and $\omega\psi$ is

$$z_1 = \tau(x_1),$$
$$z_2 = 0, \qquad x_1 \in I_1, x_2 \in I_2.$$

This map is entirely linear. Its image is a parallelotope in E^r and is therefore r-finite. Since $B_r \cap K$ may be covered by countably many neighborhoods $\phi^{-1}\sigma^{-1}(I_1 \times I_2)$, it follows that $f(B_r \cap K)$ is r-sigmafinite, as was to be shown.

COROLLARY. *If f is a C^∞-map, then fA_r is $(r+\rho)$-null for all $\rho > 0$.*

This corollary, at least in the case in which N is finite dimensional, is due to Dubovickii [3].

4. Local reduction of dimension. Consider a C^q-map θ of an open set $\circledS \subset E^{r+k} = E^r \times E^k$ into a Banach space Y, where

$$1 \leqq k < \infty, \qquad 0 \leqq r < \infty, \qquad 1 \leqq q \leqq \infty.$$

Choose an arbitrary set of axes in E^k.

LEMMA 6. *Suppose that*

$$v_0 \in E^r, \qquad w_0 = (w_0^1, \cdots, w_0^k) \in E^k, \qquad (v_0, w_0) \in \circledS,$$
$$\theta(v_0, w_0) = 0, \qquad D_2\theta(v_0, w_0) \neq 0,$$

where D_2 indicates the partial derivative relative to $w \in E^k$ when $v \in E^r$ is fixed. Then there is at least one coordinate in E^k, say the k-th, for which the following holds. There exist an open neighborhood Ω of $(v_0, w_0) \in E^{r+k}$, an open neighborhood $\bar{\Omega}$ of $(v_0, w_0^1, \cdots w_0^{k-1}) \in E^{r+k-1}$, and a C^q-map η of $\bar{\Omega}$ into E^1 such that

$$\theta(v, w) = 0, \qquad (v, w) \in \Omega, \tag{16}$$

implies that

$$w^k = \eta(v, w^1, \cdots, w^{k-1}), \qquad (v, w^1, \cdots, w^{k-1}) \in \bar{\Omega}. \tag{17}$$

The lemma does not assert that (17) implies (16). Lemma 6, adequate for our purposes, may be strengthened to allow E^r to be an arbitrary Banach space.

Proof. Suppose first that $Y = E^1$. One partial derivative of θ with respect to a component of $w \in E^k$ must be $\neq 0$. Say that $\dfrac{\partial \theta}{\partial w^k} \neq 0$. Then by the implicit function theorem, open neighborhoods Ω and $\bar{\Omega}$ and a C^q-map η exist such that (16) does indeed imply (17) and, in this case, (17) implies (16).

Now suppose that Y is an arbitrary Banach space. By hypothesis the map $D_2\theta(v_0, w_0) \neq 0$. Therefore an element w^* of E^k exists such that

$$D_2\theta(v_0, w_0)(w^* - w_0) \neq 0 \in Y.$$

Put

$$E^1 = \operatorname{span} D_2\theta(v_0, w_0)(w^* - w_0) \subset Y,$$

and

$$\theta^* = P\theta,$$

where P is a linear continuous projection onto E^1. Such projections surely exist since E^1 is one-dimensional. (See references in Part 1 of the main proof.)

Then (16) implies that

$$\theta^*(v, w) = 0.$$

Also

$$D_2\theta^*(v_0, w_0) = PD_2\theta(v_0, w_0) \neq 0,$$

since

$$PD_2\theta(v_0, w_0)(w^* - w_0) = D_2\theta(v_0, w_0)(w^* - w_0) \neq 0.$$

Now θ^* is a map into E^1. By the first paragraph, then, Ω, $\bar{\Omega}$, and η exist such that

$$\theta^*(v, w) = 0, \quad (v, w) \in \Omega,$$

implies (17). A fortiori (16) implies (17); and the proof is complete.

5. Known results; open questions. Theorem 1 is satisfactory in several ways: ρ may be any positive number; the manifold N which contains the range of the map is very general; the dimensionality of N does not enter in the condition on q. The hypothesis

$$q \geqq c(k, \rho),$$

however, is unduly demanding. It may be of interest to discuss the difficulties in the way of obtaining stronger theorems.

Suppose for the rest of this section that N is finite dimensional. Thus f is a map of an m-dimensional manifold M into an n-dimensional manifold N, $1 \leqq m, n < \infty$; M, N, and f are C^q, $1 \leqq q \leqq \infty$; M and N may include boundary; M has a countable basis; A_r denotes the set of points of M of rank $\leqq r$ relative to f. The following facts are known.

(18) fA_0 is ρ-null, if $q \geqq m/\rho$, $\rho > 0$ [22, Th. 6.1].

(19) fA_{m-1} is m-null if $q \geqq 1$ [22, Th. 4.1] and [24, Th. 1].

(20) fA_{n-1} is n-null if $q \geqq m - n + 1$, but fA_{n-1} may be n-positive if $q = m - n$ [22, Th. 7.2 and § 8].

$fA_m = fM$ is m-sigmafinite, as was established at the beginning of the proof of Theorem 1.

$fA_n = fM$ is n-sigmafinite, whether N has countable basis or not, since the image of a suitable neighborhood of each point of M is n-finite.

Knopp and Schmidt established (19) for the case $m = n$ in 1926 [12, p. 379]. The proof of (18) depends on A. P. Morse's powerful theorem on the decomposition of sets of critical points of functions [15, Th. 4.2]. The proof of (20) depends on (18) and the theorem of Fubini. Whitney's example [27] uses his powerful theorem on the extension of functions which are C^q on closed sets [26]. An alternative construction of examples is de Rham's [21]. The papers [5; 6; 7] of Dubovickii deal with theorems related to (18)-(20).

Now consider the question, partially answered by Theorem 1: When may we assert that fA_r is $(r+\rho)$-null for specified $\rho > 0$? Proceeding as in § 3, we see that it would be sufficient to show that hS is $(r+\rho)$-null, where h is the map (11) and S is the set (13). We saw in Part 1 that $(x_1, x_2) \in S$ if and only if x_2 is of rank 0 relative to the map h with x_1 fixed of an open set in E^k into E^{n-r}. By (11) the cross section of hS obtained by fixing $y_1 = \tau(x_1)$, $x_1 \in X_1$, is itself the image under h with x_1 fixed of its points of rank 0. By (18), then, the cross section of hS, y_1 fixed, is ρ-null if $q \geqq k/\rho$, for all $y_1 \in Y_1$, since the map h with x_1 fixed has a k-dimensional domain.

An important fact in the theory of Hausdorff measure is that the following generalization of Fubini's theorem is not valid.

Statement 1. Consider a compact set $B \subset E^r \times E^{n-r}$, $1 \leqq r \leqq n-1$. Let B_{y_1} denote the cross section

$$\{y_2 : (y_1, y_2) \in B\} \subset E^{n-r}$$

of B at $y_1 \in E^r$. Suppose that B_{y_1} is ρ-null for each $y_1 \in E^r$. Then B is $(r+\rho)$-null.

This statement, open until May 25, 1964, has been shown to be false by Federer, who has constructed the following counterexample. Let p be a Peano map, single valued and continuous, of the unit square in E^r onto the entire unit square in E^{n-r}. Put

$$B = \{(y_1, y_2) : y_2 = p(y_1), y_1 \in E^r\}.$$

Then B_{y_1} is, by construction, either a single point or empty. Therefore B_{y_1} is ρ-null, for any $\rho > 0$, for all $y_1 \in E^r$. Nonetheless B is $(r+\rho)$-positive if $r + \rho \leqq n - r$. For,

$$\mu_{n-r} B \geqq \mu_{n-r} \operatorname{Proj}_{E^{n-r}} B > 0,$$

since the projection of B on E^{n-r} is the unit square. Thus Statement 1 is false, at least in the case $r + \rho \leqq n - r$.

Statement 1 is true if $\rho = n - r$, by Fubini's theorem; it is true also if B_{y_1} is independent of y_1 [10, Lemma 4.1].

We have seen that the cross sections of hS, y_1 fixed, are ρ-null if $q \geqq k/\rho$. Statement 1, if it had been true, would have implied the following.

STATEMENT 2. fA_r is $(r+\rho)$-null if $q \geqq (m-r)/\rho$, $\rho > 0$.

Except for the case $r = 0$, this statement is open. Cf. (18) above. Statement 2 for $\rho = 1$ is the following.

STATEMENT 3. fA_r is $(r+1)$-null if $q \geqq m - r$.

Except for the cases $r = m-1$, $r = n-1$, and $r = 0$, this statement is open. Cf. (18), (19), (20) above.

A consequence of Statement 3 is the following theorem, which we prove directly. The theorem is due to Church who, however, assumes that $q \geqq m$ [3, Prop. 1.3].

THEOREM 2. *fA_r is of topological dimension $\leqq r$ if $q \geqq m - r$.*

Proof. It is sufficient to show that hS is of dimension $\leqq r$ [11, p. 30]. Our discussion before Statement 1 shows that each cross section of hS, y_1 fixed, is 1-null if $q \geqq (m-r)/1 = m - r$. Each such cross section then is of dimension $\leqq 0$, by the theorem of Szpilrajn (Marczewski) [11, p. 104]. Now the following analogue of Statement 1 is known: If the cross section B_{y_1} at y_1 is of dimension $\leqq 0$ for each $y_1 \in E^r$, then B is of dimension $\leqq r$ [11, p. 92]. It follows that hS is of dimension $\leqq r$, as was to be shown.

QUEENS COLLEGE,
THE CITY UNIVERSITY OF NEW YORK.

REFERENCES.

[1] A. S. Besicovitch, "A problem on measure," *Proceedings of the Cambridge Philosophical Society*, vol. 59 (1963), pp. 251-253.

[2] A. B. Brown, "Functional dependence," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 379-394.

[3] P. T. Church, "Differentiable open maps on manifolds," *Transactions of the American Mathematical Society*, vol. 109 (1963), pp. 87-100.

[4] J. Dieudonné, *Foundations of modern analysis*, Academic Press, New York, 1960.

[5] A. Ya. Dubovickii, "On differentiable mappings of an n-dimensional cube into a k-dimensional cube," (Russian), *Matematičeskii Sbornik, Novaja Serija*, vol. 32 (74) (1953), pp. 443-464.

[6] ———, "On the structure of level sets of differentiable mappings of an n-dimensional cube into a k-dimensional cube," (Russian), *Izvestija Akademii Nauk SSSR, Serija Matematičeskaja*, vol. 21 (1957), pp. 371-408.

[7] ———, "Points of complete degeneracy of the Jacobi matrix," (Russian), *ibid.*, vol. 22 (1958), pp. 705-716.

[8] ———, "Sets of points of given degeneracy of an infinitely differentiable map," (approximate title, Russian), *ibid.*, vol. 26 (1962), pp. 489-494.

[9] N. Dunford and J. Schwartz, *Linear operators, Part I*, Interscience Publishers, New York, 1958.

[10] H. Federer, "Some integralgeometric theorems," *Transactions of the American Mathematical Society*, vol. 77 (1954), pp. 238-261.

[11] W. Hurewicz and H. Wallman, *Dimension theory*, Princeton University Press, Princeton, 1941.

[12] K. Knopp and R. Schmidt, "Funktionaldeterminanten und Abhängigkeit von Funktionen," *Mathematische Zeitschrift*, vol. 25 (1926), pp. 373-381.

[13] I. Kupka, "A counter example to the theorem of Morse-Sard on infinite dimensional manifolds," to appear in *Proceedings of the American Mathematical Society*.

[14] S. Lang, *Introduction to differentiable manifolds*, Interscience Publishers, New York, 1962.

[15] A. P. Morse, "The behaviour of a function on its critical set," *Annals of Mathematics*, vol. 40 (1939), pp. 62-70.

[16] M. Morse and A. Sard, "On the measure of the critical values of functions," 1935 (unpublished).

[17] J. R. Munkres, *Elementary differential topology*, Princeton University Press, Princeton, 1963.

[18] R. S. Palais, "Morse theory on Hilbert manifolds," *Topology*, vol. 2 (1963), pp. 299-340.

[19] L. S. Pontryagin, *Smooth manifolds and their applications in homotopy theory*, (Russian), Trudy Matematičeskogo Instituta im. V. A. Steklova, no. 45, Akademii Nauk SSSR, 1955 = American Mathematical Society Translations, Series 2, vol. 11 (1959), pp. 1-114.

[20] G. de Rham, "La théorie de formes différentielles extérieures et l'homologie des variétés différentiables," *Rendiconti di Matematica e delle sue Applicazioni*, vol. 20 (1961), pp. 105-146.

[21] ———, "Sur quelques fonctions différentiables dont toutes les valeurs sont des valeurs critiques," *Celebrazioni Archimedee del Secolo XX* (Siracura, 1961), vol. II, pp. 61-65. Edizioni Oderisi, Gubbio, 1962.

[22] A. Sard, "The measure of the critical values of differentiable maps," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 883-890.

[23] ———, "The equivalence of n-measure and Lebesgue measure in E_n," *Bulletin of the American Mathematical Society*, vol. 49 (1943), pp. 758-759.

[24] ———, "Images of critical sets," *Annals of Mathematics*. vol. 68 (1958), pp. 247-259.

[25] ———, "Highly critical sets" (Abstract), *Notices of the American Mathematical Society*, vol. 10 (1963), pp. 571.

[26] H. Whitney, "Analytic extensions of differentiable functions defined in closed sets," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 63-89.

[27] ———, "A function not constant on a connected set of critical points," *Duke Mathematical Journal*, vol. 1 (1935), pp. 514-517.

ASYMPTOTIC DISTRIBUTION OF EIGENVALUES AND EIGEN-FUNCTIONS FOR NON-LOCAL ELLIPTIC BOUNDARY VALUE PROBLEMS, I.

By Felix E. Browder.*

Introduction. Let Ω be an open set in R^n, m and r positive integers, and let $W^{m,2}(\Omega)$ be the Hilbert space of r-vector functions u on Ω all of whose derivatives $D^\alpha u$ for $|\alpha| \leqq m$ lie in $L^2(\Omega)$.

In a preceding paper ([16]) the writer defined a realization of a (proper) non-local elliptic boundary value problem, for an elliptic system A of r differential operators of order m on Ω, to be a closed linear operator T in $L^2(\Omega)$ such that

$$C_c^\infty(\Omega) \subset D(T) \subset W^{m,2}(\Omega)$$

(where $D(T)$ is the domain of T, $C_c^\infty(\Omega)$ the family of C^∞-functions u with compact support in Ω) while

$$Tu = Au, \qquad u \in D(T),$$

the differential operator A being taken in the sense of the theory of distributions.

In the present paper, we are concerned with the asymptotic distribution of the eigenvalues and eigenfunctions of T. More precisely, we assume that Ω is bounded and smoothly bounded and that T is a self-adjoint operator in $L^2(\Omega)$ (i.e. $T = T^*$). Then there exists a complete orthonormal set in $L^2(\Omega)$ consisting of eigenfunctions ϕ_j of T with real eigenvalues λ_j. We consider the functions

$$c(\lambda; x, y) = \sum_{|\lambda_j| \leqq \lambda} \phi_j(x)^* \otimes \phi_j(y),$$

$$c(\lambda) = \sum_{|\lambda_j| \leqq \lambda} 1.$$

($* =$ complex conjugate), and ask for the asymptotic behaviour of $c(\lambda, x, y)$ and $c(\lambda)$ as $\lambda \to +\infty$.

We say that T is regular of order $j \geqq 0$ if $Tu \in W^{k,2}(\Omega)$ for $0 \leqq k \leqq j$ implies that $u \in W^{m+k,2}(\Omega)$. Then our basic results are the following:

Received May 8, 1964.

* Supported in part by the Sloan Foundation, National Science Foundation grants NSF-G19751 and NSF-GP2283, and Army Research Office (DURHAM) grant ARO(D)-31-124-G455.

THEOREM 1. *Suppose T is a self-adjoint operator in $L^2(\Omega)$ which is the realization of an elliptic boundary value problem for an elliptic system A of order m. Let s be the least positive integer such that $sm > n/2$, $(n = \dim \Omega)$, and suppose that the coefficients of A lie in $C^{ms-m}(\Omega)$ and that T is regular of order $m(s-1)$. Then for x and y in Ω, $(x \neq y)$*

$$\begin{cases} c(x,x,\lambda) \sim h(s,m,n)\lambda^{n/m} \int_{R^n} \{A_0(x,\eta)^{2s} + I\}^{-1} d\eta, \\ \lambda^{-n/m} c(x,y,\lambda) \to 0 \end{cases}$$

as $\lambda \to +\infty$, where $A_0(x,\eta)$ is the homogeneous characteristic matrix of the system A and

$$h(s,m,n) = (2\pi)^{-n} k(s,m,n) = (2\pi)^{-n} \left(\frac{2ms}{n}\right) \left(\frac{1}{\pi} \sin \frac{\pi n}{2ms}\right).$$

THEOREM 2. *Under the hypotheses of Theorem 1,*

$$c(\lambda) \sim h(s,m,n)\lambda^{n/m} \int_{\Omega} \operatorname{tr}\{ \int_{R^n} \{A_0(x,\eta)^{2s} + I\}^{-1} d\eta\} dx$$

as $\lambda \to +\infty$.

The proof of these two theorems is based upon the Tauberian argument of Carleman [17] combined with the asymptotic estimation of the kernel of $(T^{2s} + tI)^{-1}$ as $t \to +\infty$. The method of estimation is that indicated by the writer for single elliptic operators in the two notes [5], [8]. It is an extension of the method applied in the Dirichlet problem by L. Gårding [21] for a strongly elliptic operator A with constant coefficients and $m > n$ and was applied to the Dirichlet problem for A with variable coefficients by the writer in [5]. Using a more complicated technique of estimation based on the parametrix method, Gårding obtained the eigenvalue distribution for the Dirichlet problem for a strongly elliptic operator A with variable coefficients in [22] and the asymptotic behaviour of the spectral function (generalizing the eigenfunction distribution) for a general semi-bounded self-adjoint extension of a single strongly elliptic operator in [23]. The writer extended this latter result to semi-elliptic operators in [8] using our present type of argument. G. Ehrling in [20] applied Gårding's parametrix arguments to obtain the eigenvalue distribution for semi-bounded elliptic boundary value problems of variational type for a single strongly elliptic operator with $m > n$. J. Odhnoff [28], L. Sandgren [31], and G. Bergendal [35] have studied other eigenvalue problems and summability for eigenfunction expansions. But An Ton in his Ph. D. dissertation [33], written under the writer's direction, has obtained the eigenvalue distribution for a general class of semi-

bounded differential elliptic boundary value problems using a refined form of the parametrix argument.

The results stated in Theorems 1 and 2 are valid for any order m and for the most general non-local self-adjoint elliptic boundary value problem. The proof is both simple and direct, and uses no complicated estimation. To bring out its simplicity as sharply as possible, we restrict ourselves in the present paper to the proof of Theorems 1 and 2 and leave extensions of these results to a following paper. The extensions include: (1) boundary value problems in vector bundles; (2) non-selfadjoint problems; (3) distribution of derivatives of eigenfunctions; (4) asymptotic behaviour of the spectral function of realization whose resolvents are non-compact; (5) error estimates. All of these are treated systematically by variants of the method given below.

We should emphasize in addition that we make no semi-boundedness assumptions on the realization T whose eigenfunctions and eigenvalues we study.

In Section 1, we present our basic notation and terminology and establish some necessary auxiliary results. Section 2 gives the proofs of Theorems 1 and 2 based upon the asymptotic estimates established in Theorems 3 and 4.

Section 1. Let n be a positive integer, Ω a bounded open set in R^n. Let m be a given positive integer, s the least positive integer such that $S^m > n/2$. We shall assume throughout that the boundary of Ω is locally a $(n-1)$-dimensional C^{ms}-manifold with a C^{ms}-imbedding in R^n.

We shall consider two kinds of functions on Ω or on R^n, complex r-vector functions u for a given positive integer r (i.e. functions u from R^n or Ω to C^r) and functions U from R^n or Ω to $L(C^r, C^r)$ (which, after choice of a distinguished basis in C^r, we may think of as complex $r \times r$-matrix functions). We shall denote the first type of function by lower case letters and second by capital letters. If $u\colon R^n \to C^r$, $U\colon R^n \to L(C^r, C^r)$, Uu denotes the function $R^n \to C^r$ given by

$$(Uu)(x) = U(u(x)), \quad x \in R^n.$$

Here as usual we let $x = (x_1, \cdots, x_n)$ denote the general point of Ω or R^n, and we let dx denote the n-dimensional element of volume. Thus

$$\int_\Omega u(x)\,dx \quad \text{or} \int_{R^n} U(\eta)\,d\eta$$

denote the Lebesgue integrals of the corresponding functions.

Let $L^2(\Omega) = \{u \mid \|u\|^2 = \int_\Omega |u(x)|^2\,dx < +\infty\}$ be the usual L^2-space

with inner product

$$(u, v) = \int_\Omega \langle u(x) v(x) \rangle dx$$

where $\langle \cdot , \rangle$ is the usual inner product in C^r, $L^2(\Omega)$ the corresponding L^2-space of functions U with

$$\| U \|^2 = \int_\Omega | U(x) |^2 dx$$

where for each x in Ω, $| U(x) |$ denotes the operator norm of $U(x)$ as a linear transformation from C^r to C^r.

For each integer j, $1 \leqq j \leqq n$, we set

$$D_j = \frac{1 \partial}{i \partial x_j} \qquad (i^2 = -1)$$

while for each n-tuple of non-negative integers $\alpha = (\alpha_1, \cdots, \alpha_n)$,

$$D^\alpha = \prod_{j=1}^{n} D_j^{\alpha_j}.$$

Then $D^\alpha u = (D^\alpha u_1, \cdots, D^\alpha u_r)$ if $u = (u_1, \cdots, u_r)$, and $| \alpha | = \sum_{j=1}^{n} \alpha_j$, derivatives taken in the sense of the theory of distributions.

Definition (1.1). Let

$$W^{j,2}(\Omega) = \{ u \mid u, D^\alpha u \in L^2(\Omega) \text{ for } | \alpha | \leqq j \}, \tag{1.1}$$

$$W_I^{j,2}(\Omega) = \{ U \mid U, D^\alpha U \in L^2(\Omega) \text{ for } | \alpha | \leqq j \}, \tag{1.2}$$

$$\| u \|_j^2 = \sum_{|\alpha| \leqq j} \| D^\alpha u \|^2_{L^2(\Omega)}, \tag{1.3}$$

$$(u, v)_j = \sum_{|\alpha| \leqq j} (D^\alpha u, D^\alpha v). \tag{1.4}$$

Each $W^{j,2}(\Omega)$ and $W_I^{j,2}(\Omega)$ is a Hilbert space with respect to the corresponding inner product.

Let $C_0^\infty(\Omega)$ be the family of C^∞-functions u with compact support in Ω.

Suppose A is a linear elliptic system of r-differential operators of order m on Ω acting on r-vector functions u, i.e.

$$Au = \sum_{|\alpha| \leqq m} A_\alpha(x) D^\alpha u.$$

A is said to be elliptic if for each $x \in \Omega$ and $\xi \in R^n$, $\xi \neq 0$, the characteristic matrix

$$A(x, \xi) = \sum_{|\alpha| \leqq m} A_\alpha(x) \xi^\alpha, \quad (\xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}),$$

is non-singular.

Definition (1.2). *Let T be a closed linear operator in $L^2(\Omega)$. Then T is said to be a realization of A under an elliptic boundary condition if*

$$C_c^\infty(\Omega) \subset D(T) \subset W^{m,2}(\Omega) \tag{1.5}$$

(where $D(T)$ is the domain of T) and

$$Tu = Au, \quad u \in D(T). \tag{1.6}$$

LEMMA (1.1). *If T is the realization of A under an elliptic boundary condition, there exists a constant $c > 0$ such that*

$$\| u \|_m{}^2 \leqq c\{\| Tu \|^2 + \| u \|^2\} \tag{1.7}$$

for all u in $D(T)$.

Proof of Lemma (1.1). Consider the closed mapping of the graph of T into $W^{m,2}(\Omega)$ given by $[u, Tu] \to u$ and apply the closed graph theorem.

Definition (1.3). *The realization T of A is said to be regular of order k if $Tu \in W^{j,2}(\Omega)$ implies $u \in W^{m+j,2}(\Omega)$ for $0 \leqq j \leqq k$.*

LEMMA (1.2). *Under our hypotheses on Ω, the injection map of $W^{m,2}(\Omega)$ into $L^2(\Omega)$ is a compact linear mapping.*

Proof of Lemma (1.2). This is a part of the Sobolev Imbedding Theorem (cf. Lemma 5 of [12]).

LEMMA (1.3). *Let T be the realization of A under an elliptic boundary condition and suppose that T is self-adjoint* (i.e. $T = T^*$). *Then for each λ in the resolvent set of T, the maps $(T - \lambda I)^{-1}$ is a compact operator in $L^2(\Omega)$. There exists a complete orthonormal set $\{\phi_j\}$ in $L^2(\Omega)$ where each ϕ_j is an eigenfunction of T with real eigenvalue λ_j,*

$$T\phi_j = \lambda_j \phi_j.$$

The only limit points of the set $\{\lambda_j\}$ are $+\infty$.

Proof of Lemma (1.3). Since T is self-adjoint, its resolvent set $r(T)$ includes the complement of the real axis. For $\lambda \in r(T)$, $(T - \lambda I)^{-1}$ maps $L^2(\Omega)$ into $D(T) \subset W^{m,2}(\Omega)$. Since $(T - \lambda I)^{-1}$ is a bounded map in $L^2(\Omega)$, it is a closed map from $L^2(\Omega)$ to $W^{m,2}(\Omega)$. By the closed graph theorem, $(T - \lambda I)^{-1}$ is continuous from $L^2(\Omega)$ to $W^{m,2}(\Omega)$. Since the injection map of $W^{m,2}(\Omega)$ into $L^2(\Omega)$ is compact, $(T - \lambda I)^{-1}$ is a compact map from $L^2(\Omega)$ to $L^2(\Omega)$. The other assertions of the Lemma follow from the self-adjointness of T and the compactness of the resolvent. Q. E. D.

LEMMA (1.4). *Let Ω be a bounded open set in R^n with boundary of*

class C^{ms}. Then there exists a bounded linear map E of $L^2(\Omega)$ into $L^2(R^n)$ with

$$Eu = u \text{ on } \Omega$$

such that E maps $W^{j,2}(\Omega)$ boundedly into $W^{j,2}(R^n)$ for $0 \leqq j \leqq ms$.

Proof of Lemma (1.4). This is Proposition 2 of [4].

LEMMA (1.5). *If Ω is a bounded open set in R^n with boundary of class C^{ms} and if $ms > n/2$, then $W^{ms,2}(\Omega)$ is composed of continuous functions on Ω.*

If $\| u \|_{\infty} = \sup\{ | u(x) | : x \in \Omega\}$, there exists a constant $c_2 > 0$ such that

$$\| u \|^2_{\infty} \leqq c_2 t^{(n/2ms)-1} \{ \| u \|^2_{ms} + t \| u \|^2 \} \tag{1.8}$$

for all u in $W^{ms,2}(\Omega)$, $t \geqq 1$.

Proof of Lemma (1.5). Let $u_1 = Eu$. Then $u_1 \in W^{ms,2}(R^n)$ and if u_1 is continuous, so is u. Moreover if (1.8) holds for u_1, then

$$\begin{aligned} \| u \|^2_{\infty} \leqq \| u_1 \|^2_{\infty} &\leqq c_2 t^{(n/2ms)-1} \{ \| u_1 \|^2_{ms} + t \| u_1 \|^2 \} \\ &\leqq c_2 c t^{(n/2ms)-1} \{ \| u \|^2_{ms} + t \| u \|^2 \}. \end{aligned}$$

Hence it suffices to prove the Lemma for $\Omega = R^n$.

Since $C_c^\infty(R^n)$ is dense in $W^{ms,2}(R^n)$, it suffices to prove the validity of the inequality (1.8) for all $u \in C_c^\infty(R^n)$. Let $x \in R^n$ and let $\hat{u}$ be the Fourier transform of u, i. e.

$$\hat{u}(\xi) = (2\pi)^{-n/2} \int e^{-i\langle x,\xi\rangle} u(x)\, dx, \; \xi \in R^n.$$

Then

$$\begin{aligned} | u(x) |^2 &\leqq (2\pi)^{-n} \Big(\int_{R^n} | \hat{u}(\xi) | \, d\xi \Big)^2 \\ &\leqq (2\pi)^{-n} \Big(\int_{R^n} | \hat{u}(\xi) |^2 (| \xi |^{2ms} + t) \, d\xi \Big) \Big(\int \frac{d\xi}{| \xi |^{2ms} + t} \Big). \end{aligned}$$

Here,

$$\int | \hat{u}(\xi) |^2 (| \xi |^{2ms} + t) \, d\xi \leqq \| u \|^2_{ms} + t \| u \|^2$$

while

$$\int \frac{d\xi}{| \xi |^{2ms} + t} = t^{(n/2ms)-1} \int \frac{d\eta}{| \eta |^{2ms} + 1} = c_2 t^{(n/2ms)-1}. \qquad \text{Q. E. D.}$$

LEMMA (1.6). *Let $ms > n/2$, and let T be an elliptic realization of the system A with T self-adjoint and regular of order $m(s-1)$. Then:*

(a) *There exists a constant* $c_3 > 0$ *such that*

$$\| u \|^2_{ms} \leqq c_3 \{ \| T^s u \|^2 + \| u \|^2 \} \tag{1.9}$$

for all $u \in D(T^s)$;

(b) *There exists a constant* $c_4 > 0$ *such that for all* $t \geqq 1$,

$$\| u \|^2_{ms} + t \| u \|^2 \leqq c_4^2 \{ \| T^s u \|^2 + t \| u \|^2 \}. \tag{1.10}$$

Proof of Lemma (1.6). Since T is regular of class $m(s-1)$, $D(T^s) \subset W^{ms,2}(\Omega)$. The inequality (1.9) then follows from the closed graph theorem. The inequality (1.10) follows from inequality (1.9) by taking c_4 sufficiently large.

LEMMA (1.7). *Suppose* T *satisfies the hypotheses of Lemma* (1.6). *Then for each* $t > 0$, $(T^{2s} + tI)^{-1}$ *exists and is an operator of trace class in* $L^2(\Omega)$. *There exists a bounded measurable* $(r \times r)$*-matrix function* G_t *on* $\Omega \times \Omega$ *such that for all* f *in* $L^2(\Omega)$, x *in* Ω, *we have*

$$\{ (T^{2s} + tI)^{-1} f \}^* (x) = \int_\Omega G_t(x, y) f^*(y) \, dy.$$

For fixed x *in* Ω, $G_t(x, \cdot)$ *lies in* $W_l^{ms,2}(\Omega)$, *is continuous on* Ω, *and there exists a constant* $c_5 > 0$ *such that*

$$| G_t(x, y) | \leqq c_5 t^{(n/2ms)-1} \tag{1.11}$$

for all x *and* y *in* Ω.

Proof of Lemma (1.7). T^{2s} is a non-negative self-adjoint operator in $L^2(\Omega)$. Hence for $t > 0$, $(T^{2s} + tI)^{-1}$ exists as a bounded operator in $L^2(\Omega)$. Since T is regular of order $m(s-1)$, $D(T^s) \subset W^{ms,2}(\Omega)$. By Lemma (1.5), $W^{ms,2}(\Omega) \subset C^0(\bar{\Omega})$ and the injection mapping is bounded. Since $D(T^{2s}) \subset D(T^s)$, it follows that the range of $(T^{2s} + tI)^{-1}$ is contained in $C^0(\Omega)$. The linear mapping $(T^{2s} + tI)^{-1}$ of $L^2(\Omega)$ into $C^0(\Omega) \subset L^\infty(\Omega)$ is closed and hence bounded. We apply the Dunford-Pettis Theorem ([19], vol. 1) and obtain the existence of a unique kernel G_t on $\Omega \times \Omega$ with

$$\text{ess. sup.} \int | G_t(x, y) |^2 \, dy < +\infty$$

such that for almost all x in Ω

$$\{ (T^{2s} + tI)^{-1} f \}^* (x) = \int_\Omega G_t(x, y), f^*(y) \, dy.$$

for all f in $L^2(\Omega)$.

We must now verify that G_t satisfies the further conditions of our Lemma.

Let H_t be the Hilbert space obtained from $D(T^s)$ by imposing on it the inner product

$$(u,v)_{H_t} = (T^s u, T^s v) + t(u,v).$$

For each x in Ω, it follows from the inequalities (1.8) and (1.10) that

$$|u(x)| \leqq c_6 t^{(n/4ms)-\frac{1}{2}} \| u \|_{H_t}, \tag{1.12}$$

for all u in H_t. Hence there exists a function $G_{t,x}$ from Ω to $L(C^r, C^r)$ such that every row of $G_{t,x}$ lies in H_t, for each v in C^r

$$\| G_{t,x} v_0 \|^2_{H_t} \leqq c_7 t^{(n/2ms)-1} | v_0 |^2 \tag{1.13}$$

and for all u in H_t,

$$\langle v_0, u(x) \rangle = (G_{t,x} v_0, T^s u) + t(G_{t,x} v_0, u). \tag{1.14}$$

For u in $D(T^{2s})$

$$\langle v_0, u(x) \rangle = (G_{t,x} v_0, (T^{2s} + tI)u).$$

If $u = (T^{2s} + tI)^{-1} f$, for $f \in L^2(\Omega)$,

$$\{(T^{2s} + tI)^{-1} f\}^*(x) = \int_\Omega G_{t,x}(y) f^*(y)\, dy. \tag{1.15}$$

By the uniqueness of the kernel $G_{t,x}(y) = G_t(x,y)$ for almost all y in Ω for each given x in Ω. Hence changing $G_t(x,y)$ on a set of measure zero in $\Omega \times \Omega$, we may assume that $G_t(x,y) = G_{t,x}(y)$ and hence that $G_t(x,y)$ is continuous in y on $\bar{\Omega}$ for each x in Ω and that $G_t(x,\cdot)$ lies in H_t for each x in Ω.

Finally

$$| G_{t,x}(y) | \leqq c_8 t^{(n/2ms)-1}, \quad \| G_{t,x} v_0 \|_{H_t} \leqq c_9 t^{(n/4ms)-\frac{1}{2}} | v_0 |$$

and the proof of Lemma (1.7) is complete.

LEMMA (1.8). *If T satisfies the hypotheses of Lemma (1.7), then*

$$G_t(x,y) = \sum_j \frac{1}{\lambda_j^{2s} + t} \phi_j(x)^* \otimes \phi_j(y), \tag{1.16}$$

the series converging uniformly in y on Ω for each fixed x in Ω.

Proof of Lemma (1.8). $\{\phi_j\}$ is a complete orthonormal set in $L^2(\Omega)$. Since each ϕ_j is an eigenfunction of T, ϕ_j lies in $D(T^s)$ and hence in H_t. In addition,

$$\begin{aligned}(\phi_j, \phi_k)_{H_t} &= (T^s\phi_j, T^s\phi_k) + t(\phi_j, \phi_k) \\ &= (\lambda_j^s \lambda_k^s + t)(\phi_j, \phi_k) = (\lambda_j^s \lambda_k^s + t)\delta_j^k.\end{aligned}$$

Thus the sequence

$$\{(\lambda_j^{2s} + t)^{-\frac{1}{2}} \phi_j\}$$

is an orthonormal set in H_t. Furthermore it is complete in H_t since for any f in H_t orthogonal in H_t to all the ϕ_j, we have

$$0=(\phi_j,f)_{H_t}=(T^s\phi_j,T^sf)+t(\phi_j,f)$$
$$=((T^{2s}+t)\phi_j,f)=(\lambda_j^{2s}+t)(\phi_j,f)$$

so that $(f,\phi_j)=0$ for all j and hence $f=0$.

For fixed x in Ω, consider the Fourier expansion of $G_{t,x}$ in H_t in terms of this complete orthonormal set. We have for each v_0 in C^r,

$$(G_{t,x}v_0,(\lambda_j^{2s}+t)^{-\frac{1}{2}}\phi_j)_{H_t}=\langle v_0,(\lambda_j^{2s}+t)^{-\frac{1}{2}}\phi_j(x)\rangle.$$

Hence

$$G_{t,x}(y)=\sum_j\frac{1}{\lambda_j^{2s}+t}\,\phi_j(x)^*\otimes\phi_j(y)$$

each row of the series converging in H_t. Since convergence in H_t implies uniform convergence on Ω by Lemma (1.5), the series converges uniformly in y for fixed x in Ω, $t>0$. Q. E. D.

LEMMA (1.9). *For u in H_t, $0\leqq j<ms$, there exists a constant $c_j>0$ such that*

(1.17) $$\|u\|_j\leqq c_jt^{-(ms-j)/2ms}\|u\|_{H_t}.$$

Proof of Lemma (1.19). We know from Section 1 of [5] that

(1.18) $$\|u\|_j\leqq c_j\|u\|_{ms}{}^{j/ms}\|u\|^{(ms-j)/ms}.$$

However, $\|u\|_{ms}\leqq\|u\|_{H_t}$ while

$$\|u\|\leqq t^{-\frac{1}{2}}\|u\|_{H_t}.$$

Hence

$$\|u_j\|\leqq c_j\|u\|_{H_t}{}^{j/ms}t^{-(ms-j)/2ms}\|u\|_{H_t}{}^{(ms-j)/ms}$$
$$\leqq c_jt^{-(ms-j)/2ms}\|u\|_{H_t}. \qquad \text{Q. E. D.}$$

(The inequality of Lemma (1.9) was first published by Ehrling [20] though it was known to the writer in 1953. Another brief but more self-contained proof runs as follows: By Lemma (1.4), it suffices to prove the inequality (1.17) for u in $W^{ms,2}(R^n)$ or even in $C_c^\infty(R^n)$. We consider the inequality

$$|\xi|^{2j}\leqq c_j^2t^{-(ms-j)/ms}\{|\xi|^{2ms}+t\},\ \xi\in R^n,$$

multiply by $|\hat{u}(\xi)|^2$, and integrate in ξ over R^n.)

LEMMA (1.10). *Let $\sigma(\lambda)$ be a monotonic non-decreasing function of λ on $(0,\infty)$ and let*

$$s(t)=\int_0^\infty \frac{d\sigma(\lambda)}{\lambda+t},\ t>0.$$

Suppose $s(t)\sim ct^{a-1}$ as $t\to\infty$, $0<a<1$. Then

$$\sigma(\lambda)\sim k(a)c\lambda^a,\ \lambda\to+\infty$$

where

$$k(a)=\frac{1}{\pi a}\sin(\pi a).$$

Proof of Lemma (1.10). This is a Tauberian theorem of Hardy and Littlewood [24].

Section 2. Let A be an elliptic system of rank r and order m on $\Omega\subset R^n$, as above, $A_0(x,\eta)$ the characteristic matrix of A for $\eta\in R^n$, A' the formal adjoint of A given by

$$A'u=\sum_{|\alpha|\leq m} D^\alpha(A_\alpha(x)^*u) \tag{2.1}$$

where B^* is the adjoint of the linear transformation B in the Hilbert space C^r.

LEMMA (2.1). *If A has a self-adjoint realization, then $A'=A$ and $A_0(x,\eta)$ is a Hermitian linear transformation on C^r for $\eta\in R^n$.*

Proof of Lemma (2.1). If $T=T^*$, then for $\phi,\psi\in C_0^\infty(\Omega)$,

$$(A\phi,\psi)=(T\phi,\psi)=(\phi,T\psi)=(\phi,A\psi)$$

and

$$(A\phi,\psi)=(\phi,A'\psi).$$

Hence $A\psi=A'\psi$ for all $\psi\in C_0^\infty(R^n)$ and $A=A'$. Since

$$A'(x,\eta)=A(x,\eta)^*,$$

we see that $A(x,\eta)=A(x,\eta)^*$. Q.E.D.

Definition (2.1). *Let $x\in\Omega$, s the least positive integer such that $sm>n/2$. Then W_x is the Hilbert space given by*

$$W_x=\{f\mid f\in L^2(R^n),\ \int_{R^n}[|A_0(x,\eta)^sf(\eta)|^2+|f(\eta)|^2]d\eta<+\infty\}$$

with norm

$$\|f\|_{W_x}=\{\int_{R^n}[|A_0(x,\eta)^sf(\eta)|^2+|f(\eta)|^2]d\eta\}^{\frac{1}{2}} \tag{2.2}$$

and inner product

$$(f,g)_{W_x}=\int_{R^n}\{\langle A_0(x,\eta)^s f(\eta), A_0(x,\eta)^s g(\eta)\rangle + \langle f(\eta), g(\eta)\rangle\}d\eta. \tag{2.3}$$

The Fourier transform Fu or $\hat{u}$ of a function $u\in L^2(R^n)$ is given by

$$\hat{u}(\xi)=(Fu)(\xi)=(2\pi)^{-n/2}\int_{R^n}e^{-i\langle x,\xi\rangle}u(x)\,dx.$$

LEMMA (2.2). *$FC_0^\infty(R^n)$ is a dense subset of W_x for each x in Ω.*

Proof of Lemma (2.2). Since for $\phi\in C_0^\infty(R^n)$ $F\phi$ is a rapidly decreasing function, $FC_0^\infty(R^n)$ is a subset of W_x. Suppose it is not dense in W_x for a given x in Ω. Then there exists $g\neq 0$ in X_x such that $(\hat{\phi}, g)_W=0$ for all $\phi\in C_0^\infty(R^n)$, i.e.

$$\int_{R^n}\langle\hat{\phi}(\eta),\ (A_0(x,\eta)^{2s}+I)g(\eta)\rangle d\eta=0. \tag{2.4}$$

Replacing $\phi(x)$ by $\phi_y(x)=\phi(x-y)$ for a given y in R^n, we have

$$F(\phi_y)(\xi)=(2\pi)^{-n/2}\int_{R^n}e^{-i\langle x,\xi\rangle}\phi(x-y)\,dx$$
$$=e^{-i\langle y,\xi\rangle}\hat{\phi}(\xi).$$

Hence equation (2.4) becomes

$$0=\int_{R^n}e^{-i\langle y,\eta\rangle}\langle\hat{\phi}(\eta),\{A_0(x,\eta)^{2s}+1\}g(\eta)\rangle d\eta$$

for all y in R^n. Hence by the uniqueness theorem for Fourier transforms in L^1

$$\langle\hat{\phi}(\eta),\{A_0(x,\eta)^{2s}+I\}g(\eta)\rangle=0$$

a.e. in R^n and since $\hat{\phi}(\eta)$ fills out C^r as ϕ runs through $C_0^\infty(R^n)$,

$$\{A_0(x,\eta)^{2s}+I\}g(n)=0$$

a.e. in R^n. Since $(A_0(x,\eta)^{2s}+I\}$ is non-singular for all η in R^n, $g(\eta)=0$ a.e. Q.E.D.

LEMMA (2.3). *If $\psi\in C_0^\infty(R^n)$, and if for $\rho>0$, $x\in\Omega$,*

$$\psi_\rho(y)=\psi(x+\rho^{-1}y)$$

then

$$\hat{\psi}_\rho(\xi)=\rho^n e^{i\langle\rho x,\xi\rangle}\hat{\psi}(\rho\xi).$$

Let $N_d(x)=\{y\mid |y-x|<d\}$, $N_d(0)=\{y\mid |y|<d\}$. *Then as* ψ *runs through* $C_c^\infty(N_d(x))$, ψ_ρ *runs through* $C_c^\infty(N_{\rho d}(0))$, *and we have*

$$\bigcup_{\rho>0}\{\psi_\rho\mid\psi\in C_c^\infty(N_d(x))\}=C_c^\infty(R^n).$$

Proof of Lemma (2.3). The proof is obvious by computation.

LEMMA (2.4). *For each* v_0 *in* C^r, *if* w *is the function given by* $w(\eta)=(A_0(x,\eta)^{2s}+I)^{-1}v_0$, *then* w *lies in* W_x.

Proof of Lemma (2.4). We form the norm

$$\int_{R^n}\langle(A_0(x,\eta)^{2s}+I)w(\eta),w(\eta)\rangle d\eta$$
$$=\int_{R^n}\langle v_0,(A_0(x,\eta)^{2s}+I)^{-1}v_0\rangle d\eta$$

and remarking that by the homogeneity and ellipticity of A_0,

$$\|(A_0(x,\eta)^{2s}+I)^{-1}v_0\|\leqq\frac{c\,|v_0|}{1+|\eta|^{2sm}}$$

we see that

$$\int_{R^n}\langle v_0,(A_0(x,\eta)^{2s}+I)^{-1}v_0\rangle d\eta\leqq c\,|v_0|^2\int_{R^n}\frac{d\eta}{1+|\eta|^{2sm}}<+\infty$$

since by hypothesis $2sm>n$. Q. E. D.

THEOREM 3. *Under the hypotheses of Theorem* 1, *if* $x\in\Omega$,

$$G_t(x,x)\sim(2\pi)^{-n}t^{(n/2ms)-1}\int_{R^n}[A_0(x,\eta)+I]^{-1}d\eta \tag{2.5}$$

as $t\to+\infty$.

Proof of Theorem 3. By equation (1.14) of the proof of Lemma (1.7), for each constant vector v_0 in C^r,

$$\langle v_0,u(x)\rangle=(T^sG_{t,x}v_0,T^su)+t(G_{t,x}v_0,u) \tag{2.6}$$

for all u in $D(T^s)$, where $G_{t,x}(y)=G_t(x,y)$. For fixed x *and* t, $G_{t,x}v_0$ lies in $W^{ms,2}(\Omega)$. By Lemma (1.4), we may assume it to be the restriction of an element of $W^{ms,2}(R^n)$ which we denote once more by $G_{t,x}v_0$. Furthermore because of the inequality (1.13), we may assume that for some constant $c>0$

$$\|G_{t,x}v_0\|^2_{H_t}\leqq c^2t^{(n/2ms)-1}\,|v_0|^2. \tag{2.7}$$

It further follows from Lemma (1.10) that for $0\leqq j<ms$,

$$\begin{aligned}\|G_{t,x}v_0\|_j&\leqq c_jt^{-(ms-j)/2ms}\,\|G_{t,x}v_0\|_{H_t}\\&\leqq cc_jt^{-(ms-j)/2ms}t^{(n/4ms)-\frac{1}{2}}\,|v_0|^2.\end{aligned} \tag{2.8}$$

We choose $d_0 > 0$ such that $N_{d_0}(x)$ is contained in Ω. Then for $d < d_0$ each x in $C_c^\infty(N_d)$ is contained in $C_c^\infty(\Omega) \subset D(T^s)$ and equation (2.6) implies that

$$\langle v_0, \psi(x)\rangle = (A^s G_{t,x} v_0, A^s\psi) + t(G_{t,x}v_0, \psi). \tag{2.9}$$

Let

$$A_h = \sum_{|\alpha|=m} A_\alpha(y) D^\alpha,$$
$$A_0 = \sum_{|\alpha|=m} A_\alpha(x) D^\alpha,$$
$$R = \sum_{|\alpha|<m} A_\alpha(y) D^\alpha,$$

where the D^α are differential operators in y.

Then $A = A_h + R$, $A^s = A_h{}^s + R_1$ where R_1 is a differential operator with continuous coefficients of order less than ms. Finally

$$A^s = A_0{}^s + (A_h{}^s - A_0{}^s) + R_1 = A_0{}^s + R_2$$

where R_2 is a differential operator of order ms with continuous coefficients of the form

$$R_2 u = \sum_{|\alpha|\leqq ms} R_\alpha(x) D^\alpha u$$

where $|R_\alpha| \leqq k_0$ for all α and for $|\alpha| = ms$

$$|R_\alpha(y)| \leqq c\,|y - x|, \text{ as } y \to x.$$

Thus equation (2.9) becomes

$$\langle v_0, \psi(x)\rangle = (A_0{}^s G_{t,x} v_0, A_0{}^s\psi) + t(G_{t,x}v_0, \psi) + r_t(\psi) \tag{2.10}$$

where

$$\begin{aligned} r_t(\psi) &= ((A^s - A_0{}^s) G_{t,x} v_0, A^s\psi) + (A_0{}^s G_{t,x} v_0, (A^s - A_0{}^s)\psi) \\ &= (R_2 G_{t,x} v_0, A^s\psi) + (A_0{}^s G_{t,x} v_0, R_2\psi) \end{aligned} \tag{2.11}$$

satisfies the inequality

$$\begin{aligned} |r_t(\psi)| \leqq c\{&\| G_{t,x}v_0 \|_{ms-1} \| \psi \|_{ms} + \| G_{t,x}v_0 \|_{ms} \| \psi \|_{ms-1} \\ &+ d \| G_{t,x}v_0 \|_{ms} \| \psi \|_{ms}\} \end{aligned} \tag{2.12}$$

for $\psi \in C_c^\infty(N_d(x))$. Applying the inequality (1.17) of Lemma (1.9), we have

$$|r_t(\psi)| \leqq s\{t^{-1/2ms} \| G_{t,x}v_0 \|_{H_t} \| \psi \|_{H_t} + d \| G_{t,x}v_0 \|_{H_t} \| \psi \|_{H_t}\}. \tag{2.13}$$

Using the estimate of Lemma (1.7) on $G_{t,x}$, we have

$$|r_t(\psi)| \leqq s(t^{-1/2ms} + d)\, t^{(n/4ms - \frac{1}{2})} \| \psi \|_{H_t} \cdot |v_0| \tag{2.14}$$

for $\psi \in C_c^\infty(N_d(x))$.

Using Fourier transforms, we may rewrite equation (2.10) as

$$
\begin{aligned}
&(2\pi)^{-n/2}\int_{R^n}\langle v_0,\hat{\psi}(\xi)\rangle e^{-i\langle x,\xi\rangle}d\xi \\
(2.15)\qquad &= \int_{R^n}\{\langle A_0(x,\xi)^s F(G_{t,x}v_0)(\xi), A_0(x,\xi)^s\hat{\psi}(\xi)\rangle \\
&\quad + t\langle F(G_{t,x}v_0)(\xi),\hat{\psi}(\xi)\rangle\}d\xi + r_t(\psi).
\end{aligned}
$$

In the two integrals in equation (2.15), we make the change of variables $\xi = \rho\eta$ where $\rho = t^{1/2ms}$. Then

$$
\begin{aligned}
&\int_{R^n}\langle v_0,\hat{\psi}(\xi)\rangle e^{-i\langle x,\xi\rangle}d\xi = \rho^n\int_{R^n}\langle v_0,\hat{\psi}(\rho\eta)\rangle e^{-i\langle\rho x,\eta\rangle}d\eta \\
(2.16)\qquad &= \int_{R^n}\langle v_0, F(\psi_\rho)(\eta)\rangle d\eta = \int_{R^n}\langle [A_0(x,\eta)^{2s}+I] \\
&\qquad\qquad \times [A_0(x,\eta)^{2s}+I]^{-1}v_0, F(\psi_\rho)\rangle d\eta \\
&= (w, F(\psi_\rho))_{W_s}
\end{aligned}
$$

where ψ_ρ is the transform of ψ described in Lemma (2.3) and w is the element of W_s defined in Lemma (2.4).

Applying the same change of variables to the other integral in equation (2.15), we obtain

$$
\begin{aligned}
&\int_{R^n}\langle (A_0(x,\xi)^{2s}+tI)F(G_{t,x}v_0)(\xi),\hat{\psi}(\xi)\rangle d\xi \\
(2.17)\qquad &= t\int_{R^n}\{\langle (A_0(x,\eta)^{2s}+I)F(G_{t,x}v_0)(\rho\eta),\hat{\psi}(\rho\eta)\rangle\rho^n d\eta \\
&= (j_{t,x}, F(\psi_\rho))_{W_s}
\end{aligned}
$$

where ψ_ρ is again the function defined in Lemma (2.3) and

$$j_{t,x}(\eta) = tF(G_{t,x}v_0)(\rho\eta)e^{i\langle\rho x,\eta\rangle}$$

(where we recall that $\rho = t^{1/2ms}$).

Combining equations (2.15), (2.16), and (2.17), we have

$$(2.18)\qquad ((2\pi)^{-n/2}w - j_{t,x}, F(\psi_\rho))_{W_s} = r_t(\psi)$$

where $r_t(\psi)$ satisfies the inequality (2.14).

Let ϕ be any element of $C_0^\infty(R^n)$ with support contained in the ball $\{y \mid |y| < M\}$. Then $\phi = \psi_\rho$ for some $\psi \in C_0^\infty(N_d(x))$ if

$$\rho d = M$$

i.e. if

$$d = t^{-1/2ms}M$$

From equation (2.18) and the inequality (2.14), we then have

$$(2.19) \qquad |((2\pi)^{-n/2}w - j_{t,x}, F(\phi))_{W_x}| \leqq c_1 t^{-1/2ms} t^{(n/4ms-\frac{1}{2})} \| \psi \|_{H_t}$$

where $\psi(y) = \phi(\rho(u-x))$, and

$$\begin{aligned}
\| \psi \|^2_{H_t} &\leqq c \int_{R^n} \langle (A_0(x,\xi)^{2s} + tI)\hat{\psi}(\xi), \hat{\psi}(\xi) \rangle d\xi \\
&= c\rho^n t \int_{R^n} \langle (A_0(x,\eta)^{2s} + I)\hat{\psi}(\rho\eta), \hat{\psi}(\rho\eta) \rangle d\eta \\
&= ct\rho^{-n} \| F(\phi) \|^2_{W_x} \\
&= ct^{1-n/2ms} \| F(\phi) \|^2_{W_x}.
\end{aligned}$$

Hence

$$t^{n/4ms-\frac{1}{2}} \| \psi \|_{H_t} \leqq c_2$$

as $t \to \infty$, and the inequality (2.19) implies that

$$(2.20) \qquad |((2\pi)^{-n/2}w - j_{t,x}, F(\phi))_{W_x}| \leqq c_3 t^{-1/2ms}$$

as $t \to +\infty$.

We consider the function $t \to j_{t,x}$ from $R^+ = \{t \mid t > 0\}$ to W_x and verify that the W_x-norm of $j_{t,x}$ is bounded as $t \to +\infty$. Indeed

$$\begin{aligned}
\| j_{t,x} \|^2_{W_x} &= \int_{R^n} \langle (A_0(x,\eta)^{2s} + I) j_{t,x}(\eta), j_{t,x}(\eta) \rangle d\eta \\
&= t^2 \int_{R^n} \langle (A_0(x,\eta)^{2s} + I) F(G_{t,x}v_0)(\rho\eta), F(G_{t,x}v_0)(\rho\eta) \rangle d\eta \\
&= t\rho^{-n} \int_{R^n} \langle (A_0(x,\xi)^{2s} + tI) F(G_{t,x}v_0)(\xi), F(G_{t,x}v_0)(\xi) \rangle d\xi \\
&= t^{1-n/2ms} \| G_{t,x}v_0 \|^2_{H_t} \leqq c_4
\end{aligned}$$

by the inequality (1.13).

Since $F(C_0^\infty(R^n))$ is dense in W_x by Lemma (2.2) and $j_{t,x}$ is uniformly bounded in W_x norm and converges to $(2\pi)^{-n/2}w$ against any $F(\phi)$, $\phi \in C_0^\infty(R^n)$, it follows that $j_{t,x} \to (2\pi)^{-n/2}w$ weakly in W_x as $t \to +\infty$. Let $w_1 = (A(x,\eta)^{2s} + I)^{-1}v_1$ for a given $v_1 \in C^r$. Then by Lemma (2.4) $w_1 \in W_x$, so that

$$(2.20) \qquad (j_{t,x}, w_1)_{W_x} \to (2\pi)^{-n/2}(w, w_1)_{W_x}.$$

Since

$$(j_{t,x}, w_1)_{W_s} = \int_{R^n} \langle (A_0(x,\eta)^{2s} + I)\,(tF(G_{t,x}v_0)\,(\rho\eta)\,e^{i\langle\rho x,\eta\rangle}),$$

$$(A_0(x,\eta)^{2s} + I)^{-1}v_1\rangle d\eta$$

$$(2.21) \qquad = \int_{R^n} \langle te^{i\langle\rho x,\eta\rangle} F\,(G_{t,x}v_0)\,(\rho\eta), v_1\rangle d\eta$$

$$= t^{1-n/2ms} \langle \int_{R^n} e^{i\langle x,\xi\rangle} F(G_{t,x}v_0)\,d\xi, v_1\rangle$$

$$= (2\pi)^{n/2} t^{1-n/2ms} \langle G_{t,x}(x)\,v_0, v_1\rangle$$

while

$$(w, w_1)_{W_s} = \int_{R^n} \langle (A_0(x,\eta)^{2s} + I)\,(A_0(x,\eta)^{2s} + I)^{-1}v_0,$$

$$(A_0(x,\eta)^{2s} + I)^{-1}v_1\rangle d\eta$$

$$(2.22) \qquad = \int_{R^n} \langle (A_0(x,\eta)^{2s} + I)^{-1}v_0, v_1\rangle d\eta$$

$$= \langle \int_{R^n} (A_0(x,\eta)^{2s} + I)^{-1}v_0 d\eta, v_1\rangle,$$

it follows that

$$(2.23) \qquad t^{1-n/2ms} G_{t,x}(x) = t^{1-n/2ms} G_t(x,x) \rightarrow (2\pi)^{-n} \int_{R^n} (A_0(x,\eta)^{2s} + I)^{-1} d\eta.$$

Q. E. D.

THEOREM 4. *Under the hypotheses of Theorem* 1, *if* $x \neq y$,

$$(2.24) \qquad t^{1-n/2ms} G_t(x,y) \rightarrow 0$$

as $t \rightarrow +\infty$.

Proof of Theorem 3. If $\psi \in C_c^\infty(N_d(y))$ and $d < |x - y|$, then

$$(2.25) \qquad 0 = (T^s G_{t,x}v_0, T^s\psi) + t(G_{t,x}v_0, \psi).$$

We rewrite (2.25) as before, obtaining

$$(2.26) \qquad 0 = (G_{t,x}v_0, (A_0^{2s} + tI)\psi) + r_t(\psi).$$

We take Fourier transforms, obtain

$$0 = \int_{R^n} \langle F(G_{t,x}v_0)\,(\xi), (A_0(x,\xi)^{2s} + tI)\hat{\psi}(\xi)\rangle d\xi + r_t(\psi)$$

and make the change of variables $\xi = \rho\eta$ with $\rho = t^{1/2ms}$. We obtain

$$(2.26) \qquad -r_t(\psi) = t \int_{R^n} \langle F(G_{t,x}v_0)(\rho\eta), (A_0(x,\eta)^{2s} + I)\hat{\psi}(\rho\eta)\rangle \rho^n d\eta.$$

We let

$$\psi_\rho(y_1) = \psi(y + y_1/\rho).$$

Then

$$F(\psi_\rho)(\eta) = \rho^n \hat{\psi}(\rho\eta) e^{i\langle \rho y, \eta\rangle}$$

and equation (2.26) may be written as

$$(k_{t,x,y}, F(\psi_\rho))_{W_x} = -r_t(\psi)$$

where

$$k_{t,x,y}(\eta) = tF(G_{t,x}v_0)(\rho\eta) e^{i\langle \rho y, \eta\rangle}.$$

It follows as in the proof of Theorem 3 that

$$\| k_{t,x,y} \|_{W_x} \leqq c$$

as $t \to \infty$, while for each $\phi \in C_0^\infty(R^n)$

$$(k_{t,x,y}, F(\phi))_{W_x} \to 0.$$

Hence

$$k_{t,x,y} \to 0$$

weakly in W_x and setting $w_1 = (A_0(x,\eta)^{2s} + I)^{-1}v_1 \in W_x$, we have

$$(k_{t,x,y}, w_1)_{W_x} \to 0.$$

However,

$$\begin{aligned}
&(k_{t,x,y}, w_1)_{W_x} \\
&\quad = \int_{R^n} \langle tF(G_{t,x}v_0)(\rho\eta) e^{i\langle \rho y, \eta\rangle}, (A_0(x,\eta)^{2s} + I)(A_0(x,\eta)^{2s} + I)^{-1}v_1\rangle d\eta \\
&\quad = t\langle \int_{R^n} F(G_{t,x}v_0)(\rho\eta) e^{i\langle \rho y, \eta\rangle} d\eta, v_1\rangle \\
&\quad = t^{1-n/2ms}\langle \int_{R^n} F(G_{t,x}v_0)(\xi) e^{i\langle y, \xi\rangle} d\xi, v_1\rangle \\
&\quad = t^{1-n/2ms}(2\pi)^{n/2} G_{t,x}(y) v_0.
\end{aligned}$$

Hence

$$t^{1-n/2ms} G_t(x,y) \to 0$$

as $t \to \infty$. Q. E. D.

Proof of Theorem 1. By Lemma (1.8)

$$G_t(x,y)=\sum_j \frac{1}{\lambda_j^{2s}+t}\,\phi_j(x)^*\otimes\phi_j(y)$$

i. e.

$$\langle G_t(x,y)v_0,v_1\rangle=\sum_j \frac{1}{\lambda_j^{2s}+t}\,\langle v_0,\phi_j\rangle\langle\phi_j,v_1\rangle.$$

By Theorems 3 and 4,

$$\langle G_t(x,y)v_0,v_1\rangle t^{1-n/2ms}\to\delta_x{}^y(2\pi)^{-n}\int_{R^n}\langle(A_0(x,\eta)^{2s}+I)^{-1}v_0,v_1\rangle d\eta.$$

Let $C=\int_{R^n}(A_0(x,\eta)^{2s}+I)^{-1}d\eta$. If $x=y$, $v=v_0+\zeta v_1$, $(\zeta\in C^1)$

$$\langle G_t(x,x)v,v\rangle t^{1-n/2ms}\to(2\pi)^{-n}\langle Cv,v\rangle.$$

Applying Lemma (1.10) and rewriting

$$G_t(x,y)=\int_0^\infty\frac{d\sigma(\lambda,x,y)}{\lambda+t}$$

where

$$\sigma(\lambda,x,y)=\sum_{\lambda_j^{2s}\leqq\lambda}\phi_j(x)^*\otimes\phi_j(y),$$

we see that

$$\langle G_t(x,x)v,v\rangle=\int_0^\infty\frac{d\langle\sigma(\lambda,x,x)v,v\rangle}{\lambda+t}$$

where

$$\langle\sigma(\lambda,x,x)v,v\rangle=\sum_{\lambda_j^{2s}\leqq\lambda}|\langle\phi_j,v\rangle|^2$$

is a monotone non-decreasing function of λ. By Lemma (1.10),

$$\langle\sigma(\lambda,x,x)v,v)\rangle\sim k(s,m,n)\lambda^{n/2ms}(2\pi)^{-n}\langle Cv,v\rangle.$$

However

$$\langle\sigma(\lambda,x,x)v,v\rangle=\langle\sigma(\lambda,x,x)v_0,v_0\rangle$$
$$+|\zeta|^2\langle\sigma(\lambda,x,x)v_1,v_1\rangle+2\operatorname{Re}\zeta^*\langle\sigma(\lambda,x,x)v_0,v_1\rangle$$

while

$$\lambda^{-n/2ms}\langle\sigma(\lambda,x,x)v_0,v_0\rangle\to k(s,m,n)\,(2\pi)^{-n}\langle Cv_0,v_0\rangle$$
$$\lambda^{-n/2ms}\langle\sigma(\lambda,x,x)v_1,v_1\rangle\to k(s,m,n)\,(2\pi)^{-n}\langle Cv_1,v_1\rangle.$$

Therefore

$$\operatorname{Re}(\zeta^*\langle\sigma(\lambda,x,x)v_0,v_1\rangle)\lambda^{-n/2ms}\to k(s,m,n)\,(2\pi)^{-n}\operatorname{Re}(\zeta^*\langle Cv_0,v_1\rangle)$$

and finally

$$\lambda^{-n/2ms}\langle\sigma(\lambda,x,x)v_0,v_1\rangle\to k(s,m,n)\,(2\pi)^{-n}\langle Cv_0,v_1\rangle.$$

Since

$$c(x,y,\lambda)=\sigma(\lambda^{1/2s},x,y)$$

we obtain the conclusion

$$c(x,y,\lambda)\lambda^{-n/m}\to\delta_y{}^x k(s,m,n)(2\pi)^{-n}\int_{R^n}(A_0(x,\eta)^{2s}+I)^{-1}d\eta.$$

Q. E. D.

Proof of Theorem 2. As a special case of Theorem 3, we have

$$t^{1-n/2ms}tr\{G_t(x,x)\}=t^{1-n/2ms}\sum_j\frac{|\phi_j(x)|^2}{\lambda_j^{2s}+t}$$
$$\to(2\pi)^{-n}tr\{\int_{R^n}(A_0(x,\eta^{2s}+I)^{-1}d\eta\}$$

as $t\to+\infty$. By Lemma (1.7)

$$|t^{1-n/2ms}tr\{G_t(x,x)\}|\leqq c_0$$

for all $t\geqq 1$ and all x in Ω. We apply the Lebesgue dominated convergence theorem to integrate over Ω and obtain

$$t^{1-n/2ms}\sum_j\frac{1}{\lambda_j^{2s}+t}\to\int_\Omega(2\pi)^{-n}\int_{R^n}tr([A_0(x,\eta)^{2s}+I]^{-1})\,d\eta dx$$

Apply Lemma (1.10) once more to

$$\sigma(\lambda)=\sum_{\lambda_j^{2s}\leqq\lambda}1,$$

since we have

$$t^{1-n/2ms}\int_0^\infty\frac{d\sigma(\lambda)}{\lambda+t}\sim(2\pi)^{-n}\int_\Omega\int_{R^n}tr([A_0(x,\eta)^{2s}+I]^{-1})\,d\eta dx$$

it follows that

$$\sigma(\lambda)\sim k(s,m,n)\lambda^{n/2ms}(2\pi)^{-n}\int_\Omega\int_{R^n}tr[(A_0(x,\eta)^{2s}+I)^{-1}]d\eta dx.$$

Since

$$c(\lambda)=\sigma(\lambda^{1/2s})$$

Theorem 2 follows. Q. E. D.

INSTITUTE FOR ADVANCED STUDY
AND
UNIVERSITY OF CHICAGO.

REFERENCES.

[1] S. Agmon, "The angular distribution of eigenvalues of non-self-adjoint elliptic boundary value problems of higher order," *Conference on Partial Differential Equations and Continuum Mechanics*, The University of Wisconsin Press, 1951, pp. 9-18.

[2] ———, "On the eigenfunctions and on the eigenvalues of general elliptic boundary value problems," *Communications on Pure and Applied Mathematics*, vol. 15 (1962), pp. 119-147.

[3] S. Agmon, A. Douglis and L. Nirenberg, "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions," *Communcations on Pure and Applied Mathematics*, vol. 12 (1959), pp. 623-727.

[4] R. W. Beals, "Non-local boundary value problems for linear elliptic differential operators," *American Journal of Mathematics* (to appear).

[5] F. E. Browder, "Le problème des vibrations pour un opérateur aux derivées partielles self-adjoint et du type elliptique à coefficients variables," *Comptes Rendus de l'Academie des Sciences*, Paris, vol. 236 (1953), pp. 2140-2142.

[6] ———, "On the eigenfunctions and eigenvalues of the general linear elliptic differential operator," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 39 (1953), pp. 433-439.

[7] ———, "On the regularity properties of solutions of elliptic equations," *Communications on Pure and Applied Mathematics*, vol. 9 (1956), pp. 351-361.

[8] ———, "The asymptotic distribution of eigenfunctions and eigenvalues for semi-elliptic differential operators," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 43 (1957), pp. 270-273.

[9] ———, "La theorie spectrale des opérateurs aux derivées partielles du type elliptique," *Comptes Rendus de l'Academis des Sciences*, Paris, vol. 246 (1958), pp. 526-528.

[10] ———, "Estimates and existence theorems for elliptic boundary value problems," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 45 (1959), pp. 365-372.

[11] ———, "The spectral theory of strongly elliptic differential operators," *Proceedings of the National Academy of Sciences, U.S.A.*, vol. 45 (1959), pp. 1413-1421.

[12] ———, "On the spectral theory of elliptic operators, I," *Mathematische Annalen*, vol. 142 (1961), pp. 22-130.

[13] ———, "A priori estimates for solutions of elliptic boundary value problems," *Koninkljke Nederlandse Akademie van Wetenschappen, Proceedings (I, II)*, vol. 22 (1960), pp. 145-159, pp. 160-169; (III), vol. 23 (1961), pp. 404-410.

[14] ———, "Functional analysis and partial differential equations, II," *Mathematische Annalen*, vol. 145 (1962), pp. 81-226.

[15] ———, "Functional analysis and partial differential equations, III," (*to appear*).

[16] ———, "Non-local elliptic boundary value problems," *American Journal of Mathematics*, vol. 86 (1964), pp. 735-750.

[17] T. Carleman, "Propriétes asymptotiques des fonctions fondamentales des membranes vibrantes," *Attonde Skand. Math. Congress* (1934), pp. 34-44.

[18] ———, "Über die asymptotische Verteilung der Eigenwerte partielles Differentialgleichungen," *Berichte Verh. Sächs. Akad. Wiss. Leipzig, Mat. Nat. Kl.*, vol. 88 (1936), pp. 119-132.

[19] N. Dunford and J. T. Schwartz, *Linear operators*, vols. 1 and 2, New York, 1958, 1963.

[20] G. Ehrling, "On a type of eigenvalue problems for certain elliptic differential operators," *Mathematica Scandinavica*, vol. 2 (1954), pp. 267-285.

[21] L. Gårding, "The asymptotic distribution of the eigenvalues and eigenfunctions of a general vibration problem," *Kungl. Fysiografiska Sällskapets i Lund Handlingen*, vol. 21 (1951), pp. 1-9.

[22] ———, "On the asymptotic distribution of the eigenvalues and eigenfunctions of elliptic operators," *Mathematica Scandinavica*, vol. 1 (1953), pp. 237-255.

[23] ———, "On the asymptotic properties of the spectral function belonging to a self-adjoint semi-bounded extension of an elliptic differential operator," *Kungl. Fysiografiska Sällskapets Handlingen*, vol. 24 (1954), pp. 1-18.

[24] G. H. Hardy and J. E. Littlewood, "Notes on the theory of series (XI): On Tauberian Theorems," *Proceedings of the London Mathematical Society* (2), vol. 30 (1930), pp. 22-27.

[25] L. Hörmander, *Linear partial differential operators* (Die Grundlehren der Math. Wiss., Vol. 116), Berlin, 1963.

[26] M. V. Keldych, "On the eigenvalues and eigenfunctions of some classes of non-self-adjoint equations," *Dokladi Akademii Nauk SSSR*, vol. 77 (1951), pp. 11-14.

[27] P. Malliavin, "Un theorème taubérien avec reste pour la transformée de Stieltjes," *Comptes Rendus de l'Academie des Sciences*, Paris, vol. 255 (1962), pp. 2351-2352.

[28] J. Odhnoff, "Operators generated by differential problems with eigenvalue parameter in equation and boundary condition," *Meddelanden från Lunds Universitets Matematiska Seminarium*, vol. 14 (1959), pp. 1-80.

[29] A. Pleijel, "Propriétes asymptotiques des fonctions et valeurs periodiques de certaines problèmes de vibration," *Arkiv för Matematik, Astronomi, Fysik.* 27A, No. 13 (1940), 100 pp.

[30] A. F. Ruston, "The Fredholm theory of integral equations for operators belonging to the trace class of a general Banach space," *Proceedings of the London Mathematical Society*, vol. 53 (1957), pp. 109-124.

[31] L. Sandgren, "A vibration problem," *Meddelanden från Lunds Universitets Matematiska Seminarium*, vol. 13 (1955), pp. 1-84.

[32] M. Schechter, "General boundary value problems for elliptic differential equations," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 457-486.

[33] B. A. Ton, "Asymptotic distribution of eigenvalues and eigenfunctions for general linear elliptic boundary value problems" (*to appear*).

[34] H. Weyl, "Ramifications, old and new, of the eigenvalue problem," *Bulletin of the American Mathematical Society*, vol. 56 (1950), pp. 115-139.

[35] G. Bergendal, "Convergence and summability of eigenfunction expansions connected with elliptic differential operators," *Meddelanden från Lunds Universitets Matematiska Seminarium*, vol. 15 (1959), pp. 1-63.

AUTOMORPHIC FORMS AND POINCARE SERIES FOR INFINITELY GENERATED FUCHSIAN GROUPS.*

By Lipman Bers.

To Charles Loewner in friendship and admiration

1. Statement of results. 1. Let D be a *simply connected* domain in the extended complex plane with at least two boundary points, and G a *discrete* group of conformal self-mappings $z \to A(z)$ of D. If D is the upper half-plane U or the unit disc Δ the elements $A \in G$ are Möbius transformations and G is a *Fuchsian group* (or a Fuchsoid group in Poincaré's original terminology, since we do *not* assume G to be finitely generated). While this can be always achieved by a conformal mapping, there are some advantages in considering the seemingly more general case of an arbitrary D.

Let

(1) $$q \geqq 2$$

be a fixed integer. An *automorphic form* of weight $(-2q)$ is a holomorphic solution of the functional equation

(2) $$\phi(A(z))A'(z)^q = \phi(z) \quad \text{for } z \in D,\ A \in G.$$

We require in addition that

(3) $$\phi(z) = O(|z|^{-2q}),\ z \to \infty \ \text{ if } \infty \in D.$$

Let $\lambda_D(z)|dz|$ denote the Poincaré metric in D. The automorphic forms with

(4) $$\|\phi\|_{A_q(D,G)} = \iint_{D/G} \lambda_D(z)^{2-q} |\phi(z)|\, dxdy < \infty$$

form the Banach space $A_q(D, G)$ of *integrable* forms. The automorphic forms with

(5) $$\|\phi\|_{B_q(D,G)} = \sup \lambda_D(z)^{-q} |\phi(z)| < \infty$$

form the Banach space $B_q(D, G)$ of *bounded* forms. For $\phi \in A_q(D, G)$, $\psi \in B_q(D, G)$ the *Petersson scalar product* is defined by

(6) $$(\phi, \psi)_{q,G} = \iint_{D/G} \lambda_D(z)^{2-2q} \phi(z) \overline{\psi(z)}\, dxdy.$$

Received May 5, 1964.

* Work supported by the U. S. Army Research Office (DA-ARO-(D)-31-124-G156.

In (4) and (6) the integration is performed over an arbitrary *fundamental region* ω of G in D. This means that $\omega \subset D$ is measurable, $\operatorname{mes}(\omega - \operatorname{Int}(\omega)) = 0$, $A(z_1) \neq z_2$ for $z_1, z_2 \in \operatorname{Int}(\omega)$ and $id \neq A \in G$, and $D = \bigcup_{A \in G} A(\omega)$.

If $G = \{id\}$, we write $\boldsymbol{A}_q(D)$, $\boldsymbol{B}_q(D)$ and $(\phi,\psi)_q$ instead of $\boldsymbol{A}_q(D,G)$, $\boldsymbol{B}_q(D,G)$ and $(\phi,\psi)_{q,G}$. Clearly $\boldsymbol{A}_q(D,G) \cap \boldsymbol{A}_q(D) = \{0\}$ unless G is finite, while $\boldsymbol{B}_q(D,G)$ is always a closed linear subspace of $\boldsymbol{B}_q(D)$.

2. If $D = U$ and G has a fundamental region of finite non-Euclidean area,

$$\iint_{D/G} \lambda_D(z)^2 \, dxdy < \infty, \tag{7}$$

then $\boldsymbol{A}_q(D,G) = \boldsymbol{B}_q(D,G)$ is the finite dimensional space of so-called *cusp forms*. In the general case we have

THEOREM 1. *The Petersson product establishes an anti-isomorphism between $\boldsymbol{B}_q(D,G)$ and the dual space to $\boldsymbol{A}_q(D,G)$.*

It is trivial that, for a fixed $\psi \in \boldsymbol{B}_q(D,G)$,

$$l(\phi) = (\phi,\psi)_{q,G}$$

is a continuous linear functional on $\boldsymbol{A}_q(D,G)$, of norm $\| l \| \leqq \| \psi \|_{A_q(D,G)}$. To prove Theorem 1 we will have to show that every l can be so represented and that $\psi = 0$ whenever $(\phi,\psi)_{q,G} = 0$ for all $\phi \in \boldsymbol{A}_q(D,G)$.

3. Let $\Phi(z)$, $z \in D$, be a holomorphic function. We say that $\Theta_{q,G}\Phi$ exists if

$$(\Theta_{q,G}\Phi)(z) = \sum_{A \in G} \Phi(A(z)) A'(z)^q \tag{8}$$

where the *Poincaré series* to the right converges absolutely and uniformly on compact subsets of D. In this case $\Theta_{q,G}\Phi$ is an automorphic form of weight $(-2q)$. It is known that if (7) holds, every cusp form is a Poincaré series. In the general case we have

THEOREM 2. *$\Theta_{q,G}$ is a continuous mapping of $\boldsymbol{A}_q(D)$ onto $\boldsymbol{A}_q(D,G)$.*

Thus, for $\Phi \in \boldsymbol{A}_q(D)$, $\Theta_{q,G}\Phi$ exists and every $\phi \in \boldsymbol{A}_q(D,G)$ is of this form. If $\Phi \in \boldsymbol{B}_q(D)$, however, the series in (8) may diverge. It will certainly do so if G is infinite and $\Phi \in \boldsymbol{B}_q(D,G)$. Nevertheless we have

THEOREM 3. *Every $\psi \in \boldsymbol{B}_q(D,G)$ is of the form $\psi = \Theta_{q,G}\Psi$, $\Psi \in \boldsymbol{B}_q(D)$.*

Theorems 2 and 3 supersede the results of [2]. For the sake of completeness we shall repeat some arguments from that paper.

4. Assume now that $D = U$ (the upper half-plane). Following Eichler [5] we assign to every automorphic form ϕ of weight $(-2q)$ an element of the 1-dimensional cohomology group of G with coefficients in the additive group of polynomials in one variable of degree at most $2q-2$, the *Eichler class* of ϕ (cf. **20** below). It is known that under hypothesis (7) a cusp form is uniquely determined by its Eichler class.

THEOREM 4. *If $D = U$, G is of the first kind, and the Eichler class of $\phi \in \boldsymbol{B}_q(U, G)$ vanishes, then $\phi = 0$.*

We recall that G is said to be of the *first* or *second* kind according to whether the whole real axis is or is not contained in the closure $\wedge(G)$ of the set of real fixed points of elements of G. If G is of the second kind, $\wedge(G)$ is either a perfect nowhere dense set or contains less than three points. In the latter case G is called *elementary.*

THEOREM 5. *Let $D = U$ and let G be a non-elementary group of the second kind. The Eichler class of $\phi \in \boldsymbol{B}_q(U, G)$ vanishes if and only if ϕ is orthogonal to all forms $\Theta_{q,G}\Phi$, where $\Phi \in \boldsymbol{A}_q(U)$ is a rational function with poles in $\wedge(G)$.*

If G is of the second kind, we denote by $\boldsymbol{A}_2^{*}(U, G)$ the set of those $\phi \in \boldsymbol{A}_2(U, G)$ which are *continuous* and *real* on the real axis off $\wedge(G)$.

THEOREM 6. *Let G be as in Theorem 5. The Eichler class of $\phi \in \boldsymbol{B}_2(U, Q)$ vanishes if and only if ϕ is orthogonal to $\boldsymbol{A}_2^{*}(U, G)$.*

In Theorems 5 and 6 orthogonality is meant in the sense of the Petersson product. Theorem 4 and suitably modified forms of Theorems 5 and 6 hold also for $D = \Delta$ (the unit disc.).

5. Let D_G denote the set D from which the fixed points of elements of G distinct from the identity have been removed. The set D/G has a canonical conformal structure defined by the requirement that the projection $D \to D/G$ be a holomorphic mapping. Thus D/G and $D_G/G \subset D/G$ are Riemann surfaces. Let π_1 denote the fundamental group.

THEOREM 7. *G is finitely generated if and only if $\pi_1(D_G/G)$ is.*

The statement is trivial if G is a fixed point free in D (for then $D = D_G$ and since D is simply connected G is isomorphic to $\pi_1(D/G)$). It is "well

known". in all cases. But a direct proof has the advantage of enabling one to base the theory of finitely generated Fuchsian groups on uniformization theory to which an easy access via quasiconformal mappings is now avaiable (cf. [2]). Recently Ahlfors [1] extended Theorem 7 to Kleinian groups. Our proof of Theorem 7 is based on Theorems 4 and 6. We remark that while the proof of Theorem 4 is almost trivial, the reduction of Theorem 6 to Theorem 5 depends on a device employed by Ahlfors.

2. Preliminaries. 6. Let $f(z)$ be a conformal mapping of D. The Poincaré metric has the property that

$$\lambda_D(z)\,|\,dz\,| \text{ is a conformal invariant.} \tag{9}$$

This means that $\lambda_{f(D)}(f(z))\,|\,f'(z)\,| = \lambda_D(z)$.

For every $A \in G$ set $\hat{A} = f \circ A \circ f^{-1}$. These $\hat{A}$'s form a discrete group $\hat{G}$ of conformal self-mappings of $f(D)$. For every function $\phi(\zeta)$, $\zeta \in f(D)$, set $(f^*\phi)(z) = \phi(f(z))f'(z)^q$. Noting condition (3) we verify that f^* is an isometric linear mapping of $\boldsymbol{A}_q(f(D),\hat{G})$ onto $\boldsymbol{A}_q(D,G)$ and of $\boldsymbol{B}_q(f(D),\hat{G})$ onto $\boldsymbol{B}_q(D,G)$ which preserves the Petersson product:

$$(f^*\phi, f^*\psi)_{q,G} = (\phi,\psi)_{q,\hat{G}}.$$

One also verifies that

$$\Theta_{q,G}f^*\Phi = f^*\Theta_{q,\hat{G}}\Phi$$

where the existence of one side implies that of the other. Hence it suffices to prove Theorems 1-3 for some *fixed* domain D.

7. We have that

$$\boldsymbol{A}_q(D) \subset \boldsymbol{B}_q(D), \tag{10}$$

this injection being a *continuous* mapping.

If suffices to prove this for $D = U$ (cf. **6**) and since

$$\lambda_U(z) = |\,z - \bar{z}\,|^{-1} \tag{11}$$

the assertion follows by a standard estimate:

$$|\,\phi(z)\,| \leqq \frac{4}{\pi y^2}\iint_{|\zeta-z|<y/2} |\,\phi(\zeta)\,|\,d\xi d\eta$$

$$\leqq \frac{4}{\pi y^2}\iint_{|\zeta-z|<y/2}\left(\frac{2\eta}{y}\right)^{q-2}|\,\phi(\zeta)\,|\,d\xi d\eta \leqq \frac{2^{q+2}}{\pi(2y)^q}\iint_{\eta>0}(2\eta)^{q-2}\,|\,\phi(\zeta)\,|\,d\xi d\eta,$$

so that $\|\,\phi\,\|_{B_q(U)} \leqq 2^{q+2}\pi^{-1}\,\|\,\phi\,\|_{A_q(U)}$.

8. The *Bergman kernel function* $k_D(z,\zeta)$, $z\in D$, $\zeta\in D$ may be defined by the requirements

$$k_U(z,\zeta)=-1/\pi(z-\bar{\zeta})^2, \tag{12}$$

$$k_D(z,\zeta)\,dz\,d\bar{\zeta} \text{ is a conformal invariant,} \tag{13}$$

which means that $k_{f(D)}(f(z),f(\zeta))f'(z)\overline{f'(\zeta)}=k_D(z,\zeta)$ for every conformal mapping f of D. The kernel $k_D(z,\zeta)$ is a holomorphic function of z and $\bar{\zeta}$ and

$$k_D(\zeta,z)=\overline{k_D(z,\zeta)},\qquad \pi k_D(z,z)=\lambda_D(z)^2. \tag{14}$$

Also,

$$\iint_D \lambda_D(\zeta)^{2-q}\,|\,k_D(z,\zeta)\,|^q\,d\xi d\eta=C_q\lambda_D(z)^q \tag{15}$$

where C_q is a constant. In view of (9) and (13) it suffices to verify this for $D=U$ in which case (15) follows from the identity:

$$\iint_{\eta>0}\frac{y^q\eta^{q-2}d\xi d\eta}{|\,x+iy-\xi+i\eta\,|^{2q}}=\int_{-\infty}^{+\infty}\frac{d\xi}{(1+\xi^2)^q}\int_0^{+\infty}\frac{\eta^{q-2}d\eta}{(1+\eta)^{2q-1}}$$

for $y>0$.

From now on we omit the subscript D. The following *reproducing formula* holds (as it does also in bounded homogeneous domains in several variables, cf. Selberg [6]):

$$\phi(z)=c_q\iint_D\lambda(\zeta)^{2-2q}k(z,\zeta)^q\phi(\zeta)\,d\xi d\eta \tag{16}$$

for $\phi\in\boldsymbol{B}_q(D)$, where

$$c_q=(2q-1)\pi^{q-1} \tag{16'}$$

It suffices to verify this for $D=\Delta$. Since

$$k_\Delta(z,\zeta)=1/\pi(1-z\bar{\zeta})^2,\ \lambda_\Delta(z)=(1-|\,z\,|^2)^{-1} \tag{17}$$

and

$$(2q-1)\iint_{|\zeta|<1}\frac{(1-|\,\zeta\,|^2)^{2q-2}\bar{\zeta}^m d\xi d\eta}{(1-z\bar{\zeta})^{2q}}=\pi z^m,\qquad m=0,1,\cdots$$

the assertion follows.

9. Let $L_1(D)$ and $L_\infty(D)$ denote the usual complex Banach spaces of (equivalence classes of) integrable and bounded measurable functions in D. For $\mu\in L_1(D)$ set

$$(\alpha_q\mu)(z)=c_q\iint_D\lambda(\zeta)^{-q}k(z,\zeta)^q\mu(\zeta)\,d\xi d\eta. \tag{18}$$

and for $\nu \in L_\infty(D)$ set

$$(\beta_q \nu)(z) = c_q \iint_D \lambda(\zeta)^{2-q} k(z,\zeta)^q \nu(\zeta)\, d\xi d\eta. \tag{19}$$

By (15) the mappings α_q and β_q are continuous linear mappings of $L_1(D)$ and $L_\infty(D)$ into $\boldsymbol{A}_q(D)$ and $\boldsymbol{B}_q(D)$, respectively. These mappings are *onto,* since

$$\alpha_q(\lambda^{2-q}\phi) = \phi \text{ for } \phi \in \boldsymbol{A}_q(D), \tag{20}$$

$$\beta_q(\lambda^{-q}\phi) = \phi \text{ for } \phi \in \boldsymbol{B}_q(D), \tag{21}$$

by (10) and (16). Also

$$(\alpha_q \mu, \psi)_q = \iint_D \mu(z)\lambda(z)^{-q}\overline{\psi(z)}\, dxdy \text{ for } \psi \in \boldsymbol{B}_q(D), \tag{22}$$

and

$$(\phi, \beta_q \nu)_q = \iint_D \lambda(z)^{2-q}\phi(z)\overline{\nu(z)}\, dxdy \text{ for } \phi \in \boldsymbol{A}_q(D). \tag{23}$$

The proof involves merely substitution into the definition (6) for $G = \{id\}$, a change of order of integration, and an application of (16).

10. Let l be a continuous linear functional on $\boldsymbol{A}_q(D)$. By the theorems of Hahn-Banach and F. Riesz there is a $\nu \in L_\infty(D)$ such that

$$l(\phi) = \iint_D \lambda_D(z)^{2-q}\phi(z)\overline{\nu(z)}\, dxdy.$$

Hence, by (23) we have that $l(\phi) = (\phi, \psi)_q$ where $\psi = \beta_q \nu$. Next, let $\psi \in \boldsymbol{B}_q(D)$ be such that $(\phi, \psi)_q = 0$ for all $\phi \in \boldsymbol{A}_q(D)$. Noting (22) we conclude that

$$\iint_D \lambda(z)^{-q}\overline{\psi(z)}\mu(z)\, dxdy = 0$$

for all $\mu \in L_1(D)$. Hence $\psi \equiv 0$. Thus we have proved Theorem 1 for the case $G = \{id\}$.

3. Poincaré series and duality. **11.** We prove now that $\Theta_q = \Theta_{q,G}$ *is a continuous mapping of* $\boldsymbol{A}_q(D)$ *into* $\boldsymbol{A}_q(D, G)$.

Let $\Phi \in A_q(D)$ and let ω denote a fundamental region of G in D. Then

$$\begin{aligned}\iint_\omega & \lambda(z)^{2-q} \Big| \sum_{A \in G} \Phi(A(z))A'(z)^q \Big| \, dxdy \\ &\leqq \sum_{A \in G} \iint_\omega \lambda(z)^{2-q} | \Phi(A(z))A'(z)^q | \, dxdy \\ &= \sum_{A \in G} \iint_\omega \lambda(A(z))^{2-q} | \Phi(A(z)) | \, | A'(z) |^2 \, dxdy \\ &= \sum_{A \in G} \iint_{A(\omega)} \lambda(z)^{2-q} | \Phi(z) | \, dxdy = \| \Phi \|_{A_q(D)}.\end{aligned}$$

This implies the absolute and uniform convergence of the series (8) in every compact subset of a fundamental region and hence on every compact subset of D, as well as the inequality

$$\| \Theta_q \Phi \|_{A_q(D,G)} \leqq \| \Phi \|_{A_q(D)}.$$

(Here we used two well known facts: L_1 convergence of holomorphic functions implies normal convergence. If $D_0 \subset \subset D$ there is an $\omega_0 \subset \subset \omega$ and a finite sequence $\{A_1, \cdots A_n\} \subset G$ such that $D_0 \subset A_1(\omega_0) \cup \cdots \cup A_n(\omega_0)$.)

12. Let l be a continuous linear functional on $A_q(D, G)$. Let ω be a fundamental region. Then, by Hahn-Banach and F. Riesz,

$$l(\phi) = \iint_\omega \lambda(z)^{2-q} \phi(z) \nu(z) \, dxdy \tag{24}$$

with a bounded measurable $\nu(z)$. We extend ν over the whole of D by the relation

$$\nu(A(z)) (\overline{A'(z)}/A'(z))^{q/2} = \nu(z) \quad \text{for } A \in G \tag{25}$$

(where $(\bar{A}'/A')^{q/2} = |A'|^q (A')^{-q}$). For $\Phi \in A_q(D)$ we have

$$l(\Theta_q \Phi) = \iint_D \lambda(z)^{2-q} \Phi(z) \nu(z) \, dxdy \tag{26}$$

as follows from the identity

$$\begin{aligned}\iint_\omega & \lambda(z)^{2-q} \sum_{A \in G} \Phi(A(z)) A'(z)^q \nu(z) \, dxdy \\ &= \sum_{A \in G} \iint_{A(\omega)} \lambda(z)^{2-q} \Phi(z) \nu(z) \, dxdy.\end{aligned}$$

Using this we shall show that

$$(27) \qquad l(\Theta_q \Phi) = 0 \text{ for all } \Phi \in A_q(D)$$

implies that

$$(28) \qquad l(\phi) = 0 \text{ for all } \phi \in A_q(D, G),$$

which means that

$$(29) \qquad \Theta_q A_q(D) \text{ is dense in } A_q(D, G).$$

During this proof we assume that $D = \Delta$ (the unit disc) and 0 is not a fixed point of any element of G distinct from the identity. This assumption involves no loss of generality.

13. For ν satisfying (25) for $D = \Delta$ and such that the corresponding functional l vanishes on $\Theta_q A_q(\Delta)$ set

$$(30) \qquad h(z) = -\frac{1}{\pi} \iint_{|\zeta|<1} \frac{(1-|\zeta|^2)^{q-2}\nu(\zeta)\, d\xi d\eta}{\zeta - z}$$

and, for some fixed θ, $0 < \theta < 2\pi$,

$$(31) \qquad \tilde{h}(z) = -\frac{(1-ze^{-i\theta})^{q-2}}{\pi} \iint_{|\zeta|<1} \frac{(1-|\zeta|^2)^{q-2}\nu(\zeta)\, d\xi d\eta}{(1-\zeta e^{-i\theta})^{q-2}(\zeta - z)}.$$

For a fixed z such that $|z| \geqq 1$ the functions

$$\Omega(\zeta) = -\frac{1}{\pi}\,\frac{1}{\zeta - z}, \qquad \tilde{\Omega}(\zeta) = -\frac{1}{\pi}\,\frac{1}{(1-\zeta e^{-i\theta})^{q-2}(\zeta - z)}$$

belong to $A_q(\Delta)$, and by (26)

$$h(z) = l(\Theta_q \Omega), \qquad \tilde{h}(z) = (1-ze^{-i\theta})^{q-2} l(\Theta_q \tilde{\Omega}),$$

so that by (27)

$$(32) \qquad h(z) = \tilde{h}(z) = 0 \text{ for } |z| \geqq 1.$$

From well known properties of logarithmic potentials we conclude that h and $\tilde{h}$ are continuous everywhere and that, in view of the second equation (32)

$$(33) \qquad \left| \iint_{|\zeta|<1} \frac{(1-|\zeta|^2)^{q-2}\nu(\zeta)\, d\xi d\eta}{(1-\zeta e^{-i\theta})^{q-2}(\zeta - z)} \right| \leqq c(1-|z|)\log\frac{1}{1-|z|}$$

for $|z| < 1$, where c does not depend on θ. Also

$$(34) \qquad \frac{\partial h}{\partial \bar{z}} = \frac{\partial \tilde{h}}{\partial \bar{z}} = (1-|z|^2)^{q-2}\nu(z) \text{ for } |z| < 1$$

(in the sense of weak derivatives). By (32) and (34) we have that $h \equiv \bar{h}$. Noting (33) and the fact that θ was arbitrary we conclude that

$$h(z) = O(-(1-|z|)^{q-1}\log(1-|z|)), \qquad |z|\uparrow 1. \tag{35}$$

One computes easily from (34) and (25) that for every fixed $A \in G$ the function

$$h(A(z))A'(z)^{1-q} - h(z)$$

is holomorphic in $|z| < 1$. Since it vanishes on $|z| = 1$ we have that

$$h(A(z)) = h(z)A'(z)^{q-1} \text{ for } A \in G. \tag{36}$$

Using these properties of h we shall show that $l \equiv 0$.

14. Let ω be the closure in Δ of the set

$$\{z \in \Delta \mid |A(z)| > |z| \text{ for } id \neq A \in G\}$$

and let ω_r be the intersection of ω with $|z| < r < 1$. Then ω is a fundamental region. For every r, $0 < r < 1$, the boundary σ_r of ω_r consists of a portion γ_r of the circle $|z| = r$ and of $2n = 2n(r)$ circular arcs $\delta_1, \cdots, \delta_n$, $\delta_1', \cdots, \delta_n'$ such that there exist elements $A_1, \cdots, A_n$ of G with

$$A_j(\delta_j) = -\delta_j', \qquad j = 1, \cdots, n. \tag{37}$$

All this is known and easy to check.

Now let $\phi \in A_q(\Delta, G)$ be given. By (24) and (34)

$$l(\phi) = \lim_{r\uparrow 1} \iint_{\omega_r} (1-|z|^2)^{q-2}\nu(z)\phi(z)\,dxdy$$

$$= \lim_{r\uparrow 1} \iint_{\omega_r} \phi\frac{\partial h}{\partial \bar{z}}\,dxdy = -(i/2)\lim_{r\uparrow 1}\int_{\sigma_r} \phi h\,dz.$$

Since by (34) we have that

$$\phi(A(z))h(A(z))A'(z) = \phi(z)h(z) \text{ for } A \in G,$$

it follows from (37) that

$$\int_{\delta_j} \phi h dz + \int_{\delta_j'} \phi h dz = 0, \qquad j = 1, \cdots, n,$$

so that

$$2il(\phi) = \lim_{r\uparrow 1} \int_{\gamma_r} \phi h dz,$$

and by (35)

(38) $$|l(\phi)| \leqq \text{const.} \liminf_{r \uparrow 1} (1-r)^{q-1} \log \frac{1}{1-r} \int_{\gamma_r} |\phi| \, |dz|.$$

Since

$$\int_{\frac{1}{2}}^{1} (1-r^2)^{q-2} \int_{\gamma_r} |\phi| \, |dz| \, dr \leqq \| \phi \|_{A_q(\Delta, G)}$$

(38) implies that $l(\phi) = 0$. Q. E. D.

15. For $\Phi \in \boldsymbol{A}_q(D)$, $\psi \in \boldsymbol{B}_q(D, G)$ we have that

(39) $$(\Phi, \psi)_q = (\Theta_q \Phi, \psi)_{q, G}.$$

Indeed this means that

$$\sum_{A \in G} \iint_{A(\omega)} \lambda(z)^{2-2q} \Phi(z) \overline{\psi(z)} \, dx dy$$
$$= \iint_{\omega} \lambda(z)^{2-2q} \overline{\psi(z)} \sum_{A \in G} \Phi(A(z)) A'(z)^q \, dx dy$$

which is easily verified.

16. *Proof of Theorem* 1. Assume that $\psi \in \boldsymbol{B}_q(D, G)$ is such that $(\phi, \psi)_{q, G} = 0$ for all $\phi \in \boldsymbol{A}_q(D, G)$, then $(\Theta_q \Phi, \psi)_{q, G} = 0$ for all $\Phi \in \boldsymbol{A}_q(D)$ and by (39) also $(\Phi, \psi)_q = 0$. Hence $\psi = 0$ by the result in **10.**

Now let $l(\phi)$ be a given linear functional $\boldsymbol{A}_q(D, G)$. Then (cf. **12**) there is a $\nu \in L_\infty(D)$ satisfying (25) such that (24) holds. Set $\psi = \beta_q \bar{\nu}$. Then (cf. **9**) $\psi \in \mathbf{B}_q(D)$ and by (26) and (23)

(40) $$l(\Theta_q \Phi) = (\Phi, \psi)_q \text{ for } \Phi \in \boldsymbol{A}_q.$$

Now, for $A \in G$ and $B = A^{-1}$, we have, noting (9), (13) and (25),

$$\psi(A(z)) A'(z)^q = c_q \iint_D A'(z)^q \lambda(\zeta)^{2-q} k(A(z), \zeta)^q \overline{\nu(\zeta)} \, d\xi d\eta$$
$$= c_q \iint_D \lambda(A \circ B(\zeta))^{2-q} A'(z)^q k(A(z), A \circ B(\zeta))^q \overline{\nu(A \circ B(\zeta))} \, d\xi d\eta$$
$$= c_q \iint_D \lambda(B(\zeta))^{2-q} k(z, B(\zeta))^q \overline{\nu(B(\zeta))} \, |B'(\zeta)|^2 \, d\xi d\eta = \psi(z).$$

Thus $\psi \in \boldsymbol{B}_q(D, G)$ and, by (39) and (40),

$$l(\phi) = (\phi, \psi)_{q, G}$$

whenever $\phi \in \Theta_q \boldsymbol{A}_q(D)$. In view of (29) the same holds for all $\phi \in \boldsymbol{A}_q(D, G)$.

17. *Proof of Theorem* 2. In view of **11** we must show only that $\Theta_q \boldsymbol{A}_q(D) = \boldsymbol{A}_q(D, G)$. Let $\phi \in \boldsymbol{A}_q(D, G)$ and let χ be the characteristic function of a fundamental region ω. Then $\chi\lambda^{2-q}\phi \in L_1(D)$ and we may form $\hat{\phi} = \Theta_q \alpha_q(\chi\lambda^{2-q}\phi)$ which belongs to $\boldsymbol{A}_q(D, G)$. Let ψ be any element in $\hat{\boldsymbol{B}}_q(D, G)$. By (39)

$$
\begin{aligned}
(\hat{\phi}, \psi)_{q,G} &= (\alpha_q(\chi\lambda^{2-q}\phi), \psi)_q \\
&= c_q \iint_D \lambda(z)^{2-2q}\overline{\psi(z)} \iint_\omega \lambda(\zeta)^{2-2q}\phi(\zeta) k(z, \zeta)^q d\xi d\eta dx dy \\
&= c_q \iint_\omega \lambda(\zeta)^{2-2q}\phi(\zeta) \iint_D \lambda(z)^{2-2q}\overline{k((\zeta, z)^q \psi(z)}\, dx dy d\xi d\eta \\
&= (\phi, \psi)_{q,G}.
\end{aligned}
$$

Hence

$$\phi = \Theta_q \alpha_q(\chi\lambda^{2-q}\phi), \tag{41}$$

by Theorem 1.

18. *Proof of Theorem* 3. Let χ be as in the previous proof. We shall show that if $\phi \in \boldsymbol{B}_q(D, G)$, then

$$\phi = \Theta_q \beta_q(\chi\lambda^{-q}\phi) \tag{42}$$

(note that $\chi\lambda^{-q}\phi \in L_\infty(D)$). By (16)

$$\phi(z) = \sum_{A \in G} c_q \iint_{A(\omega)} \lambda(\zeta)^{2-2q} k(z, \zeta)^q \phi(\zeta)\, d\xi d\eta$$

this series being absolutely and normally convergent. Setting $B = A^{-1}$ and using (2), (9) and (12) we obtain

$$
\begin{aligned}
\phi(z) &= \sum_{A \in G} c_q \iint_{A(\omega)} \lambda(B(\zeta))^{2-2q} k(B(z), B(\zeta))^q \phi(B(\zeta)) B'(z)^q \,|B'(\zeta)|^2\, d\xi d\eta \\
&= \sum_{A \in G} c_q B'(z)^q \iint_\omega \lambda(\zeta)^{2-2q} k(B(z), \zeta)^q \phi(\zeta) d\xi d\eta
\end{aligned}
$$

which is precisely (42).

4. Periods of automorphic forms. **19.** Let $D = U$ so that G is a group of Möbius transformations $z \to A(z) = (az + b)/(cz + d)$. Let Π_{2q-2}

denote the additive groups of polynomials $P(z)=\sum_{j=0}^{2q-2}\alpha_j z^j$. The group G operates from the right on Π_{2q-2} by the rule

$$(PA)(z)=P(A(z))A'(z)^{1-q}. \tag{43}$$

A mapping $A\rightarrow P_A$ of G into Π_{2q-2} is called a *cocycle* if

$$P_{AB}=P_AB+P_B, \tag{44}$$

a *coboundary* if there exists an element $Q\in\Pi_{2q-2}$ such that

$$P_A=QA-Q. \tag{45}$$

The coboundaries form a subgroup of the additive group of cocycles. The factor group (cocycles/coboundaries) is denoted by $H^1(G,\Pi_{2q-2})$.

20. Let ϕ be an automorphic form of weight $(-2q)$ and F a holomorphic function such that

$$\frac{d^{2q-1}F(z)}{dz^{2q-1}}=\phi(z). \tag{46}$$

One verifies easily that for every $A\in G$ the $(2q-1)$-st derivative of

$$F(A(z))A'(z)^{1-q}-F(z) \tag{47}$$

vanishes, so that this function belongs to Π_{2q-2}. We call it the *Eichler period* of F on A. The mapping

$$A\rightarrow F(A(z))A'(z)^{1-q}-F(z) \tag{48}$$

is clearly a cocycle. Since F is determined by ϕ modulo a polynomial of degree at most $2q-2$, the cohomology class of (48) depends only on ϕ and depends on ϕ linearly. We call it the *Eichler class* of ϕ.

The existence of an F satisfying (46) and the condition

$$F(A(z))A'(z)^{1-q}=F(z) \text{ for all } A\in G \tag{49}$$

is necessary and sufficient for the vanishing of the Eichler class of ϕ.

21. Let $\phi\in\boldsymbol{B}_q(U,G)$. Then $|\phi(x+iy)|\leqq\text{const.}\,y^{-q}$, so that every $F(z)$ satisfying (46) is continuous on the real axis. Assume that (49) holds and let $x\in\boldsymbol{R}$ be a fixed point of a hyperbolic or parabolic element A of G. Then $A(x)=x$, $A'(x)\neq1$ and, by (49), $F(x)=0$. Hence also

$$F(x)=0 \text{ for } x\in\wedge(G) \tag{50}$$

where $\wedge(G)$ is the *closure of the set of real fixed points*. Conversely, if (46)

and (50) hold, then for every fixed $A \in G$ the polynomial (47) vanishes on $\wedge(G)$ since $A(\wedge(G)) = \wedge(G)$ for every $A \in G$. If G is not elementary, $\wedge(G)$ is infinite and we conclude that (49) holds.

22. *Proof of Theorem* 4. If G is of the first kind and the Eichler class of $\phi \in \boldsymbol{B}_q(U, G)$ is zero, then $\phi(z) = F^{(2q-1)}(z)$ with $F = 0$ on $\boldsymbol{R}$. Hence $F \equiv 0$, $\phi \equiv 0$.

23. Let $\wedge$ be a perfect set on the real axis (in the next paragraph we shall take $\wedge = \wedge(G)$ for a non-elementary group G of the second kind). Let $a_1, \cdots, a_q$ be distinct points of $\wedge$ and set

$$p(z) = (z - a_1)(z - a_2) \cdots (z - a_q). \tag{51}$$

Then every rational function with simple poles in $\wedge$ which belongs to $\boldsymbol{A}_q(U)$ is of the form

$$\sum_{j=1}^{n} \frac{\alpha_j}{(z - x_j)p(z)} \tag{52}$$

where $x_1, \cdots, x_n$ are distinct points of $\wedge$, $x_j \neq a_k$, and the α_j are arbitrary complex constants. Indeed, a rational function with no singularities except perhaps for simple poles at $a_1, \cdots, a_q$, $x_1, \cdots, x_n$ belongs to $\boldsymbol{A}_q(U)$ if and only if it is of the form

$$\sum_{j=1}^{q} \frac{\beta_j}{z - a_j} + \sum_{j=1}^{n} \frac{\gamma_j}{z - x_j}$$

with

$$\sum_{j=1}^{q} \beta_j a_j^s + \sum_{j=1}^{n} \gamma_j x_j^s = 0, \qquad s = 0, 1, \cdots, q-1.$$

The space of such functions has therefore dimension n. On the other hand (52) always belongs to $\boldsymbol{A}_q(U)$.

If $\Phi(z) \in \boldsymbol{A}_q(U)$ is a rational function with poles in $\wedge$, it is a limit of functions of the form (52). Indeed, if $\xi_1, \cdots, \xi_m$ are the poles of Φ and $\nu_1, \cdots, \nu_m$ their multiplicities we have $0 < \nu_j \leqq q - 1$ and

$$\Phi(z) = r(z) \prod_{j=1}^{m} (z - \xi_j)^{-\nu_j}$$

where $r(z)$ is a polynomial of degree at most $\nu_1 + \cdots + \nu_m + q - 1$ with $r(\xi_j) \neq 0$. Let $\epsilon > 0$ be given. Since $\wedge$ is perfect there exist distinct points ξ_{jk}, $j = 1, \cdots, m$, $k = 1, \cdots, \nu_j$ in $\wedge$ with $|\xi_{jk} - \xi_j| < \epsilon$. The function

$$\tilde{\Phi}(z) = r(z) \prod_{j=1}^{m} \prod_{k=1}^{\nu_j} (z - \xi_{jk})^{-1}$$

is of the form (52) and one verifies that $\| \Phi - \tilde{\Phi} \|_{A_q(U)}$ will be arbitrarily small for ϵ sufficiently small.

24. *Proof of Theorem* 5. Let $\Lambda = \Lambda(G)$ and let $a_1, \cdots, a_q$ and $p(z)$ be as in **23**. Let $\phi \in \boldsymbol{B}_q(U, G)$. Noting (11), (12) we write (16) in the form

$$\phi(z) = \frac{(-1)^q(2q-1)}{\pi} \iint_{\eta>0} \frac{|\zeta - \bar{\zeta}|^{2q-2}\phi(\zeta)\, d\xi d\eta}{(\bar{\zeta} - z)^{2q}}.$$

Set

$$G(z) = \iint_{\eta>0} \frac{|\zeta - \bar{\zeta}|^{2q-2}\phi(\zeta)\, d\xi d\eta}{(\bar{\zeta} - z)\, p(\bar{\zeta})}.$$

This function is holomorphic in U and continuous everywhere except perhaps at the points a_j. Next, set

$$F(z) = \frac{(-1)^q p(z)\, G(z)}{\pi(2q-2)!}.$$

Then $F(a_j) = 0$, $j = 1, \cdots, q$ and since

$$\frac{p(z)}{p(\bar{\zeta})(\bar{\zeta} - z)} - \frac{1}{\bar{\zeta} - z}$$

is a polynomial of degree $q-1$ in z, we have that

$$F^{(2q-1)}(z) = \frac{(-1)^q(2q-1)}{\pi} \iint_{\eta>0} \frac{|\zeta - \bar{\zeta}|^{2q-2}\phi(\zeta)\, d\xi d\eta}{(\bar{\zeta} - z)^{2q}}$$

in U, so that (46) holds. By **21** the Eichler class of ϕ vanishes if and only if $F(x) = 0$ for $x \in \Lambda(G)$, $x \neq a_j$. This condition is equivalent to

$$G(x) = 0 \text{ on } \Lambda(G) - \{a_1, \cdots, a_q\}.$$

But for a real $x \neq a_j$

$$\overline{G(x)} = (\Phi, \phi)_q$$

where

$$\Phi(z) = \frac{1}{(z-x)p(z)} \in A_q(U).$$

The conclusion of Theorem 5 now follows from **23** and **15**.

25. Let G be again a Fuchsian non-elementary group of the second kind and let Ω denote the complement of $\Lambda(G)$ in the extended complex plane. Then there exists a Fuchsian group H_0 without elliptic elements and

a holomorphic mapping $\zeta \to g(\zeta)$ of U onto Ω such that if $\zeta_1, \zeta_2 \in U$, then $g(\zeta_1) = g(\zeta_2)$ if and only if there is a $C \in H_0$ with $C(\zeta_1) = \zeta_2$. Also, there is a Fuchsian group H such that if $\zeta_1, \zeta_2 \in U$, then $A(g(\zeta_1)) = g(\zeta_2)$ for some $A \in G$ if and only if there is a $B \in H$ with $B(\zeta_1) = \zeta_2$. The mapping τ of H onto G which sends $B \in H$ into $A \in G$ with $g \circ B = A \circ g$ is a homomorphism; its kernel is precisely H_0.

Let $\phi \in A_2^*(U, G)$. This means that $\phi \in A_2(U, G)$ and $\phi(z)$ is holomorphic in Ω and satisfies the relation

$$\phi(\bar{z}) = \overline{\phi(z)}. \tag{53}$$

Let ω be a fundamental region for G in Ω chosen so that $\omega \cap U$ is simply connected and ω is invariant under the mapping $z \to \bar{z}$. Then there is a fundamental region $\hat{\omega}$ for H in U such that $g(\hat{\omega}) = \omega$. Let $K \subset H$ contain exactly one representative of each coset of H modulo H_0. Then

$$\hat{\omega}_0 = \bigcup_{B \in K} B(\hat{\omega})$$

is a fundamental region for H_0 in U and $g(\hat{\omega}_0) = \Omega$.

Set $\hat{\phi}(\zeta) = \phi(g(\zeta))g'(\zeta)^2$. Then

$$\iint_{\hat{\omega}} |\hat{\phi}(\zeta)|\, d\xi d\eta = \iint_{\omega} |\phi(z)|\, dxdy = 2 \,\|\phi\|_{A_2(U,G)}$$

by (53), and for $B \in H$ we have that

$$\begin{aligned}\hat{\phi}(B(\zeta))B'(\zeta)^2 &= \phi(g(B(\zeta))g'(B(\zeta))^2B'(\zeta)^2 \\ &= \phi(A(g(\zeta))A'(g(\zeta))^2g'(\zeta)^2 = \phi(g(\zeta))g'(\zeta)^2 = \hat{\phi}(\zeta)\end{aligned}$$

where A is the image of B under the homomorphism τ described above. Hence $\hat{\phi} \in A_2(U, H)$ and by Theorem 2 we have that $\hat{\phi} = \Theta_{2,H}\hat{\Phi}$, $\hat{\Phi} \in A_2(U)$, or

$$\begin{aligned}\hat{\phi}(\zeta) &= \sum_{B \in H} \hat{\Phi}(B(\zeta))B'(\zeta)^2 \\ &= \sum_{B \in K} \sum_{C \in H_0} \hat{\Phi}(C(B(\zeta))C'(B(\zeta))^2B'(\zeta)^2.\end{aligned} \tag{54}$$

Set

$$\hat{\Phi}_0(\zeta) = \sum_{C \in H_0} \hat{\Phi}(C(\zeta))C'(\zeta)^2.$$

Then $\hat{\Phi}_0 = \Theta_{2,H_0}\hat{\Phi} \in A_2(U, H_0)$. Hence there exists a holomorphic function $\Phi_0(z)$, $z \in \Omega$ such that $\hat{\Phi}_0(\zeta) = \Phi_0(g(\zeta))g'(\zeta)^2$; we have that

$$\iint_{\Omega} |\Phi_0(z)|\, dxdy < \infty$$

since this integral equals $\|\hat{\Phi}_0\|_{A_2(U,H_0)}$.

Now (54) may be written as

$$\begin{aligned}\phi(g(\zeta))g'(\zeta)^2 &= \sum_{B\in K} \hat{\Phi}_0(B(\zeta))B'(\zeta)^2\\ &= \sum_{B\in K} \Phi_0(g(B(\zeta)))g'(B(\zeta))^2B'(\zeta)^2\\ &= \sum_{A\in G} \Phi_0(A(g(\zeta)))A'(g(\zeta))^2g'(\zeta)^2.\end{aligned}$$

Thus *every* $\phi \in A_2^\#(U,G)$ *admits the representation*

$$\phi = \Theta_{2,G}\Phi_0 \tag{55}$$

where Φ_0 *is holomorphic in* Ω *and absolutely integrable over this domain.*

26. *Proof of Theorem* 6. Assume that $\psi \in B_2(U,G)$ is orthogonal to $A_2^\#(U,G)$. Let $r(z)$ be a rational function with poles in $\wedge(G)$ belonging to $A_2(U)$ and $\phi = \Theta_{2,G}r$. Since

$$\int\int_\Omega |r(z)|\,dxdy < \infty$$

(Ω having the same meaning as in **25**) the argument in **11** can be repeated to show that the Poincaré series

$$\sum_{A\in G} r(A(z))A'(z)^2$$

converges absolutely and normally in Ω. This implies that $\phi(z) = \phi_1(z) + i\phi_2(z)$, with $\phi_1, \phi_2 \in A_2^\#(U,G)$. Hence $(\phi,\psi)_{2,G} = 0$. By Theorem 5 the Eichler class of ψ is zero.

Assume next that the Eichler class of $\psi \in B_2(U,G)$ vanishes and let $\phi \in A_2^\#(U,G)$. Then ϕ admits the representation (55). By the approximation theorem proved in [2] there exists a sequence of rational functions $\{r_j(z)\}$, with poles in $\wedge(G)$ such that

$$\int\int_\Omega |r_j(z) - \Phi_0(z)|\,dxdy \to 0. \tag{56}$$

By Theorem 5 we have that $(\Theta_{2,G}r_j,\psi)_{2,G} = 0$. Since (56) implies that $\|r_j - \Phi_0\|_{A_2(U)} \to 0$, we have that $\Theta_{2,G}r_j \to \phi$ in $A_2(U,G)$, by Theorem 2. Therefore $(\phi,\psi)_{2,G} = 0$.

5. Finitely generated Fuchsian groups. 27. A Riemann surface S will be called of *finite type*, more precisely of type (g,n,m), if it is conformally equivalent to $S_0 - \sigma$ where S_0 is a closed (compact) surface of genus g and σ a closed set with $n + m \geqq 0$ components of which $n \geqq 0$ are

points and $m \geqq 0$ simply connected non-degenerate continua. The numbers g, n, m depend only on S; we say that S has n punctures and m boundary curves.

If $m = 0$ then S_0 (the natural compactification of S) is determined by S except for conformal equivalence. If $m > 0$ there exists a Riemann surface S_1 of type $(2g+m-1, 2n)$ (the *doubling* of S) which is determined by S except for conformal equivalence, m disjoint simple closed analytic curves $\gamma_1, \cdots, \gamma_m$ on S_1 and an anticonformal involution ρ of S_1 which leaves a point $p \in S_1$ fixed if and only if $p \in \gamma = \gamma_1 \cup \cdots \cup \gamma_m$, such that $S_1 - \gamma$ consists of two components one of which is conformally equivalent to S.

The fundamental group $\pi_1(S)$ is finitely generated if and only if S is of finite type. This is a known result in surface topology.

28. Let D_G/G be of finite type. Then G is finitely generated.

This is well known and can be proved by dissecting D_G/G by finitely many smooth curves into a simply connected region such that a component of its inverse image under the projection $D_G \to D_G/G$ is a fundamental domain whose boundary consists of finitely many "sides."

29. Let S be a Riemann surface. An *Abelian differential* (of the first kind) on S is a rule associating with every local parameter $p \to t(p)$ defined on a domain $K \subset S$ a holomorphic function $\phi(t)$ such that $\phi(t)\,dt$ is invariant under parameter changes. In this case $|\phi(t)|^2$ is a density. If we demand instead the invariance of $\phi(t)\,dt^2$ we obtain a *quadratic* (holomorphic) differential; now $|\phi(t)|$ is a density. The Abelian differentials α with

$$\iint_S |\alpha|^2 < \infty$$

form a Hilbert space $A_1(S)$ of square integrable differentials. The quadratic differentials β with

$$\iint_S |\beta| < \infty \tag{57}$$

form the Banach space $A_2(S)$ of integrable differentials. We have that

$$\dim A_1(S) \leqq \dim A_2(S)$$

because if $\alpha_1, \alpha_2 \in A_1(S)$, then $\alpha_1\alpha_2 \in A_2(S)$. If the genus of S is infinite, then $\dim A_1(S) = \infty$ (cf. Nevanlinna [4]) and hence $\dim A_2(S) = \infty$. If the genus of S is $g < \infty$, then $S = S_0 - \sigma$ where σ is a closed set on the closed

surface S_0 of genus g. If S contains N distinct points, then $\dim \boldsymbol{A}_2(S) \geqq N$ since it is known (say from the Riemann-Roch theorem) that to every $p \in S_0$ there is a meromorphic quadratic differential β_p on S_0 whose only singularity is a simple pole at p. We conclude that

(58) $\qquad \dim \boldsymbol{A}_2(S) = \infty$ unless S is of finite type $(g, n, 0)$.

30. The space $\boldsymbol{A}_2(D, G)$ can be defined even when G is a discrete group of conformal self-mappings of a non-simply connected domain (since λ does not enter in the definition of this space). Let D_G denote D with the fixed points of elements of G (distinct from the identity) removed. Then there is a canonical isomorphism

$$\boldsymbol{A}_2(D, G) \cong \boldsymbol{A}_2(D_G/G). \tag{59}$$

Indeed, $\boldsymbol{A}_2(D, G)$ may be identified with the space $\boldsymbol{X}$ of meromorphic quadratic differentials β on the Riemann surface D/G for which (57) holds and which have no singularities except perhaps simple poles on the set σ consisting of the images of fixed points of G under the projection $D \rightarrow D/G$. Since σ is discrete and $D/G - \sigma = D_G/G$, $\boldsymbol{X}$ may be identified with $\boldsymbol{A}_2(D_G/G)$.

31. Let G be a Fuchsian group. The elements of $\boldsymbol{B}_q(U, G)$ with vanishing Eichler class form a closed linear subspace $\boldsymbol{B}_q{}^0(U, G)$.

If G is finitely generated, $\dim \boldsymbol{B}_q(U, G)/\boldsymbol{B}_q{}^0(U, G) < \infty$.

Indeed, assign to every $\phi \in \boldsymbol{B}_q(U, G)$ a holomorphic function $F(z)$, $z \in U$ such that $F^{(2q-1)}(z) = \phi(z)$ and $F^{(\nu)}(i) = 0$, $\nu = 0, 1, \cdots, 2q-2$. Then $\phi \in \boldsymbol{B}_q{}^0(U, G)$ whenever the Eichler periods of F vanish on a set of generators of G. This amounts to finitely many linear conditions.

32. *Proof of Theorem* 7. We may assume that $D = U$. We may assume that G is non-elementary, the theorem being trivial for elementary groups. In view of **27**, **28** it suffices to assume that G is finitely generated and to prove that U_G/G is of finite type.

Let G be of the first kind. Then $\boldsymbol{B}_2{}^0(U, G) = \{0\}$ by Theorem 4, hence $\dim \boldsymbol{B}_2(U, G) < \infty$ by **31**, hence $\dim \boldsymbol{A}_2(U, G) < \infty$ by Theorem 1, hence $\dim \boldsymbol{A}_2(U_G/G) < \infty$ by (59), hence U_G/G is of the finite type $(g, n, 0)$ by (58).

Assume next that G is of the second kind. Let $\boldsymbol{A}_2{}^\flat(U, G)$ denote the subspace of $\boldsymbol{A}_2(U, G)$ consisting of elements of the form $\phi_1 + i\phi_2$ with $\phi_1, \phi_2 \in \boldsymbol{A}_2{}^\#(U, G)$. By Theorems 1 and 6 the dual space to $\boldsymbol{A}_2{}^\flat(U, G)$ is anti-isomorphic to $\boldsymbol{B}_2(U, G)/\boldsymbol{B}_2{}^0(U, G)$. Thus $\dim \boldsymbol{A}_2{}^\flat(U, G) < \infty$. Let Ω

have the same meaning as in **25**. One sees at once that $A_2{}^b(U,G)$ may be identified with $A_2(\Omega,G)$. Hence $\dim A_2(\Omega_G/G)<\infty$ by (59), and in view of (58) the Riemann surface $S_1=\Omega_G/G$ is of finite type $(g,n,0)$. The mapping $z\rightarrow\bar{z}$ induces an anti-conformal involution ρ on S_1. The set γ of fixed points of ρ is the image of the intersection of Ω_G with the extended rela axis under the canonical mapping $\Omega_G\rightarrow S_1$ and one of the two components of $S_1-\gamma$ is U_G/G. Hence U_G/G is of finite type (g,n,m) with $m>0$.

NEW YORK UNIVERSITY.

REFERENCES.

[1] L. V. Ahlfors, "Finitely generated Kleinian groups," *American Journal of Mathematics*, vol. 86 (1964), pp. 413-429.

[2] L. Bers, "Completeness theorems for Poincaré series in one variable," *Proceedings of the International Symposium on Linear Spaces*. Jerusalem Academic Press-Pergamon Press, 1961, pp. 88-100.

[3] L. Bers, "An approximation theorem," to appear in the *Journal d'Analyse Math.* (Jerusalem).

[4] R. Nevanlinna, *Uniformisierung*. Springer, 1953.

[5] M. Eichler, "Eine Verallgemeinerung der Abelschen Integrale," *Mathematische Zeitschrift*, vol. 67 (1957), pp. 267-278.

[6] A. Selberg, "Automorphic functions and integral operators," *Seminars on Analytic Functions*, Institute for Advanced Study, 1957, vol. II, pp. 152-161.

A NOTE ON STIEFEL MANIFOLDS.

By E. H. Brown, Jr.[1] and B. Steer.

In [1] Kervaire considered a certain collection of homotopy spheres Σ^{2n-1}, where n is odd and $n \neq 1, 3, 7$, and showed that Σ^9 was not diffeomorphic to S^9. (Σ^{2n-1} is the generator of bP_{2n} [2].) In this note we show that if V_n is the Stiefel manifold of unit tangent vectors to S^n, the connected sum $V_n \# \Sigma^{2n-1}$ and V_n are diffeomorphic. We also obtain some properties of stably parallelizable manifolds with Arf invariant 1 [1].

We will say that a manifold M is of type (K) if it has the following properties. $\dim M = 2n$ where n is odd and $n \neq 1, 3, 7$. M is an $(n-1)$-connected, compact closed combinitorial manifold with a compatible differentiable structure, except possibly at one point. $M - pt$ is stably parallelizable. $H_n(M; Z) \approx Z \oplus Z$ and if $S_1{}^n$ and $S_2{}^n \subset M$ are n-spheres which represent generators of $H_n(M; z)$, then the normal bundles of $S_1{}^n$ and $S_2{}^n$ are non trivial. (It is an easy consequence of [1] that this later property does not depend on the choice of generators, i. e. M has Arf invariant 1.)

It follows from the results of [2] that any stably parallelizable manifold in the dimensions under consideration, is f-cobordant to a homotopy sphere or to a manifold of type (K). For each odd $n \neq 1, 3, 7$, a manifold of type (K) was constructed in [2] and it was shown that Σ^{2n-1} above is diffeomorphic to $\partial(M$-open disc) if M is of type (K).

Let W_n be the manifold of tangent vectors to S^n of length less than or equal to 1. Then $\partial W_n = V_n$.

THEOREM 1. *If M is of type (K), M is combinatorially equivalent to $W_n \cup_f W_n$ where $f\colon V_n \to V_n$ is a piecewise linear homeomorphism. Furthermore if M has a differentiable structure, M is diffeomorphic to $W_n \cup_f W_n$ where f is a diffeomorphism.*

THEOREM 2. *V_n and $\Sigma^{2n-1} \# V_n$ are diffeomorphic.*

Let $i\colon D^n \to S^n$ be a diffeomorphism of the n disc into S^n and let $\phi\colon D^n \times S^{n-1} \to V^n$ be a bundle map covering i. Let $\theta\colon S^{n-1} \to SO(n)$ be a

Received June 8, 1964.

[1] The results of this paper were obtained while the first named author was on a National Science Foundation Fellowship.

characteristic map for $\tau(S^n)$, the tangent bundle of S^n. Let $g: \phi(D^n \times S^{n-1}) \to \bar{\phi}(D^n \times S^{n-1})$ be defined by

$$g(\phi(x,y)) = \phi(\theta(y)x,y)$$

We choose θ so that g is differentiable.

THEOREM 3. *For a given n, there is a differentiable manifold of type (K) (and hence $\Sigma^{2n-1} = S^{2n-1}$) if and only if g can be extended to a diffeomorphism of V_n onto V_n.*

Proof of Theorems 1 *and* 2. Suppose M is a manifold of type (K), $p \in M$, and we are given a differentiable structure on $M - p$. By assumption, the generators of $H_n(M;Z)$ may be represented by spheres $S_1{}^n$, $S_2{}^n \subset M - p$ with non-trivial normal bundles. Since $M - p$ is a stably parallelizable manifold, these normal bundles are equivalent to $\tau(S^n)$. Therefore $S_1{}^n$ has a tubular neighborhood T diffeomorphic to W_n. Let $T \subset M$ be the union of T and a smooth thickened arc going from ∂T to p. If M has a differentiable structure T is also diffeomorphic to W_n. Let $N = \overline{M - T} \subset M$. Then N is a differentiable manifold and $\partial N = \Sigma^{2n-1} \# V_n$. To prove Theorems 1 and 2 it is sufficient to show that N is diffeomorphic to W_n. By Lefschetz duality $H_*(N;Z) \approx H_*(W_n;Z)$. Let $S_3{}^n \subset$ Interior N be a generator of $H_n(N;Z)$. (In M, $S_3{}^n$ and $S_1{}^n$ are isotopic.) $S_3{}^n$ has a tabular neighborhood T_3 diffeomorphic to W_n. Let $P = \overline{N - T_3}$. Then $\partial P = V_n \cup \Sigma^{2n-1} \# V_n$. By duality (or simply by an exact sequence argument) one sees that the inclusions of each of the components of ∂P into P induce isomorphisms in homology. Everything in sight is simply connected and hence P is an h-cobordism. Therefore by [3], P is diffeomorphic to $V_n \times I$ and thus N is diffeomorphic to W_n.

Proof of Theorem 3. Suppose $f: V_n \to V_n$ is a diffeomorphism. Note $S_1{}^{n-1} = \phi(\{0\} \times S^{n-1}) \subset V_n$ is a generator of $\pi_{n-1}(V_n)$. Hence f may be modified by and isotopy so that it leaves $S_1{}^{n-1}$ pointwise fixed and

$$g(\phi(x,y)) = \phi(u(y)x,y) \quad (x,y) \in D^n \times S^{n-1}$$

for some $u: S^{n-1} \to SO(u)$. Consider $M = W_n \cup_f W_n$. In M, $S_1{}^{n-1}$ bounds a disc on each side of V_n, i. e. the fibre in W_n whose boundary is $S_1{}^{n-1}$. Let $S_1{}^n$ be the union of these two discs and let $S_2{}^n$ be the zero cross section of W_n

on one side M. One may easily check that M is of type (K) if and only if the normal bundle of S_1^n is equivalent to $\tau(S^n)$. But u is a characteristic map for this bundle. This together with Theorem 1 proves Theorem 3.

BRANDEIS UNIVERSITY,
HERTFORD COLLEGE, OXFORD.

REFERENCES.

[1] M. A. Kervaire, "A manifold which does not admit any differentiable structure," *Commentarii Mathematici Helvetici*, vol. 34 (1960), pp. 257-270.

[2] M. A. Kervaire and J. W. Milnor, "Groups of homotopy spheres," *Annals of Mathematics*, vol. 77 (1963), pp. 504-537.

[3] S. Smale, "On the structure of manifolds," *American Journal of Mathematics*, vol. 84 (1963), pp. 387-399.

PARTITIONS WITH ODD SUMMANDS—SOME COMMENTS AND CORRECTIONS.

By Peter Hagis, Jr.

Theorem 2 of [2] states that the sum

$$A = \sum_{h \bmod k}' \omega(h,k) \exp\{-2\pi i(hn - h'\nu)/k\}$$

where $h \equiv d (\bmod 4)$, $2 \nmid d$; $\sigma_1 \leqq h' < \sigma_2 \ (\bmod k)$, $0 \leqq \sigma_1 < \sigma_2 \leqq k$, $2 \mid k$, is subject to the estimate $O(n^{1/3}k^{2/3+\epsilon})$ uniformly in ν, d, σ_1, σ_2. Here σ_1, σ_2 are integers, $hh' \equiv -1 \ (\bmod k)$ and $\sum'$ indicates that h runs through a reduced residue system module k. $\omega(h,k)$ is a root of unity whose exact value is given in (2.1) of [2]. It is stated in [2] that the proof of this result is identical with that of Theorem 2 of [1]. This is true if $4 \mid k$. It is *not* true if $k \equiv 2 \ (\bmod 4)$, and I have been unable to obtain an alternative proof of the theorem in this case. The purpose of this note is to prove the weaker theorem in which the restriction $h \equiv d \ (\bmod 4)$ is dropped in case $k \equiv 2 \ (\bmod 4)$. The use of this weaker result still permits one to prove the results concerning partitions with odd summands obtained in the final section of [2].

Theorem. *The sum*

$$A = \sum_{h \bmod k}' \omega(h,k) \exp\{-2\pi i(hn - h'\nu)/k\}$$

where $\sigma_1 \leqq h' < \sigma_2 \ (\bmod k)$, $0 \leqq \sigma_1 < \sigma_2 \leqq k$, $k \equiv 2 \ (\bmod 4)$ *is subject to the estimate* $O(n^{1/3}k^{2/3+\epsilon})$ *uniformly in* ν, σ_1, σ_2.

Proof. According to (4.11) in [2], if $k \equiv 2 \ (\bmod 4)$,

$$\omega(h,k) = \exp\{2\pi i(\phi(h'-h)/Gk^* \\ + \Gamma(-12d(k+k^*) + 4d + 3(2k-k^2))/f)\}.$$

Here, if $3 \mid k$ then $f = 16$, $G = 3$ while if $3 \nmid k$ then $f = 48$, $G = 1$. ϕ and Γ are defined by the congruences $f\phi \equiv 1 \ (\bmod Gk^*)$, $Gk^*\Gamma \equiv 1 \ (\bmod f)$ where $k = 2k^*$. $hh' \equiv -1 \ (\bmod Gk)$, and $d = \pm 3$ according as $h \equiv \pm 3 \ (\bmod Gk)$. Writing $F(d,k) = -12d(k+k^*) + 4d + 3(2k-k^2)$ we easily verify that

Received August 19, 1964.

$F(-3,k)-F(3,k)=72k+24(3k^*-1)$. Since $2\mid k$ and $2\mid(3k^*-1)$ we see that $F(-3,k)\equiv F(3,k)\pmod f$. Therefore,

$$\omega(h,k)=\exp\{2\pi i(\phi(h'-h)/Gk^*+C)\}$$

where C depends on k but *not* on h.

From (2.1) in [2] we see immediately that $\omega(h,k)$ has period k when viewed as a function of h. It follows from these remarks that

$$A=c(k)\sum_{h \bmod Gk}{}' \exp\{2\pi i g(h)/Gk\}$$

where $c(k)=G^{-1}\exp\{2\pi iC\}$ and $g(h)=(2\phi+G\nu)h'-(2\phi+Gn)h$.

If we now define the function $m(s)$ for all integers s by

$$m(s)=\begin{cases}1, & \text{if } \sigma_1\leqq s<\sigma_2 \pmod k,\\ 0, & \text{otherwise,}\end{cases}$$

then $m(s)$ has period k, and from the theory of finite Fourier series we have

$$m(s)=\sum_{j=0}^{k-1}\lambda_j\exp\{2\pi isj/k\} \quad\text{where}\quad \lambda_j=k^{-1}\sum_{s=0}^{k-1}m(s)\exp\{-2\pi isj/k\}.$$

We can now write

$$\begin{aligned}A&=c(k)\sum_{h \bmod Gk}{}' m(h')\exp\{2\pi ig(h)/Gk\}\\ &=c(k)\sum_{j=0}^{k-1}\lambda_j\sum_{h \bmod Gk}{}'\exp\{2\pi i((-2\phi-Gn)h+(2\phi+G\nu+Gj)h')/Gk\}\\ &=c(k)\sum_{j=0}^{k-1}\lambda_j S(k,j,n,\nu)\end{aligned}$$

where the restriction $\sigma_1\leqq h'<\sigma_2\pmod k$ has been dropped. $S(k,j,n,\nu)$ is a complete Kloosterman sum and by [5] we have

$$|S(k,j,n,\nu)|<K(Gk)^{2/3+\epsilon}(-2\phi-Gn,Gk)^{1/3}$$

where K does not depend on k, j, n or ν.

It is obvious that $(-2\phi-Gn,Gk)\leqq 2G(-2\phi-Gn,k^*)$, and since $(f,k^*)=1$, $-2f\phi=-2+Jk^*$, and $fG=48$ we have

$$(-2\phi-Gn,Gk)\leqq 2G(-2-48n,k^*)=O(n).$$

Therefore, $S(k,j,n,\nu)=O(k^{2/3+\epsilon}n^{1/3})$ and since $\sum_{j=0}^{k-1}\lambda_j=O(\log k)$ (see Section

2 in [4]), we have $A = O(k^{2/3+\epsilon}n^{1/3})$. This completes the proof of the theorem.

Remarks similar to those made in the first paragraph of this paper apply also to Theorem 3 in [3]. Thus, all references to d in this theorem should be deleted. The weakening of this theorem has no effect on the argument in Section 4 of [3]. The proof of the modified theorem is the same as that just given.

In closing I wish to take this opportunity to correct another misstatement in [3]. With $x' = \exp\{2\pi ih'/k - 2\pi/zk\}$, $2hg' \equiv -1 \pmod{k^*}$, $k = 2k^*$ it is stated just after (2.8) in [3] that $x' = \exp\{2\pi ig'/k^* - \pi/zk^*\}$. This should be $x' = -\exp\{2\pi ig'/k^* - \pi/zk^*\}$ so that in the second line of Theorem 1 of [3], $F^2(x')$ should be replaced by $F(x')F(-x')$. This change has no effect on the main results of [3].

TEMPLE UNIVERSITY, PHILADELPHIA, PENNSYLVANIA.

REFERENCES.

[1] P. Hagis, Jr., "A problem on partitions with a prime modulus $p \geqq 3$," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 30-62.

[2] ———, "Partitions into odd summands," *American Journal of Mathematics*, vol. 85 (1963), pp. 213-222.

[3] ———, "Partitions into odd and unequal parts," *American Journal of Mathematics*, vol. 86 (1964), pp. 317-324.

[4] S. Iseki, "A partition function with some congruence condition," *American Journal of Mathematics*, vol. 81 (1959), pp. 939-961.

[5] H. Salié, "Zur Abschätzung der Fourierkoeffizienten ganzer Modulformen," *Mathematische Zeitschrift*, vol. 36 (1933), pp. 263-278.

A NOTE ON THE NAKAI-MOISEZON TEST FOR AMPLENESS OF A DIVISOR.

By Steven Kleiman.*

Let V be a nonsingular variety defined over an algebraically closed field. Call a divisor D on V *arithmetically positive* if the intersection number $(D^i \cdot Y)$ is strictly positive for all i-dimensional subvarieties Y of V, $i = 1, \cdots, r$, where r is the dimension of V. Call D *ample* if some multiple nD of D is linearly equivalent to a hyperplane section for some projective embedding of V, i. e., if the rational map of V defined by $|nD|$ is biregular. Clearly if D is ample, it is arithmetically positive. Conversely, the Nakai-Moisezon test states, *if D is arithmetically positive, it is ample.*

This test was first discovered by Nakai [4] for nonsingular surfaces. Then Moisezon outlined a proof for higher dimensional nonsingular varieties in [2], and in [3] he suggested a definition of the intersection number $(D^i \cdot Y)$ on singular varieties and remarked that with that definition the test continues to be valid. Independently, Nakai [5] extended the test to projective algebraic schemes.

In Section 1 below, we give a proof of the test, which, in our opinion, is simpler and more self-contained although based on that of Nakai [5]. Since we have stated the test for nonsingular varieties, we cannot immediately apply induction on the dimension r of V. Therefore, we prove the following slightly stronger theorem, in substance stated by Moisezon [2], from which the test follows immediately by taking $U = V$:

Theorem 1. *Let U be a nonsingular projective variety, let V be an r-dimensional subvariety, and let D be a divisor on U. Suppose the intersection number $(D^i \cdot Y)$ is strictly positive for all i-dimensional subvarieties Y of V, $i = 1, \cdots, r$. Then the intersection cycle $D \cdot V$, considered as a Cartier divisor on V, is ample for V.*

Our proof of Theorem 1 easily generalizes. In Section 2, we discuss the modification necessary to prove the following theorem:

Theorem 2. *Let V be a complete algebraic scheme, and let D be a Cartier divisor on V. Suppose D is arithmetically positive. Then D is ample.*

Received July 10, 1964.

* NSF Graduate Fellow.

Here, with Nakai [5], we call D *arithmetically positive* if the Euler-Poincaré characteristic $\chi(\mathcal{O}_V(nD)\otimes\mathcal{O}_Y)$ $(=[(D^i\cdot Y)/r!]n^r+\cdots$, if V is non-singular of dimension $r)$ $\to\infty$ as $n\to\infty$ for every closed integral subscheme Y of positive dimension i.

I take this opportunity to thank Professors O. Zariski and D. Mumford for the stimulating discussions which led to the publication of this note.

1. Proof of Theorem 1. If D is a Cartier divisor on a variety V, we shall write for simplicity $\mathcal{O}(D)$ for $\mathcal{O}_V(D)$, the sheaf of germs of functions f on V such that $(f)+D>0$ locally, $H^i(D)$ for $H^i(\mathcal{O}_V(D))$, and $h^i(D)$ for $\dim H^i(\mathcal{O}_V(D))$.

Suppose we have U, V, and D given as in Theorem 1. Let $D_0=D\cdot V$ considered as a Cartier divisor on V. Our main difficulty is in showing that the complete linear system $|nD_0|$ has no base points if n is sufficiently large. First we prove that $|nD_0|$ is not empty if n is large.

Let C_2 be a general hypersurface section of V of large degree, and let C_1 be a general member of $|D_0+C_2|$. Then D_0 is linearly equivalent to C_1-C_2. Let $d_1=D\cdot C_1$ and $d_2=D\cdot C_2$, considered as Cartier divisors on C_1 and C_2, if V is r-dimensional, $r>1$. If V is a curve, C_i is a set of m_i distinct points where $m_i=\deg C_i$, $i=1,2$. Then set $\mathcal{O}(nd_i)=\mathcal{O}_{C_i}$, $i=1,2$. Consider the exact sequence of sheaves

$$\begin{array}{ccccccccc} 0 & \to & \mathcal{O}(nD_0-C_1) & \to & \mathcal{O}(nD_0) & \to & \mathcal{O}(nd_1) & \to & 0 \\ & & \| s & & & & & & \\ 0 & \to & \mathcal{O}((n-1)D_0-C_0) & \to & \mathcal{O}((n-1)D_0) & \to & \mathcal{O}((n-1)d_2) & \to & 0. \end{array}$$

Set $\chi'(nD_0)=h^0(nD_0)-h^1(nD_0)$.

If V is a curve, $h^0(\mathcal{O}(nd_1))=m_i$ and $h^1(\mathcal{O}(nd_1))=0$ for $i=1,2$ and for any n. Hence

$$\begin{aligned}\chi'(nD_0)=\chi'((n-1)D_0)&=h^0(\mathcal{O}(nd_1))-h^0(\mathcal{O}(n-1)d_2)\\ &=m_1-m_2=(D\cdot V).\end{aligned}$$

Therefore $\chi'(nD_0)=(D\cdot V)n+k$, for some constant k. Finally, since $(D\cdot V)>0$, $\chi'(nD_0)\to\infty$ as $n\to\infty$. Hence also $\dim|nD_0|=h^0(nD_0)-1\to\infty$ as $n\to\infty$. Therefore $|nD_0|$ is not empty for large n.

If V is r-dimensional, $r>1$, then by induction we may assume that d_i is ample and that

$$\begin{aligned}&h^0(nd_i)=\chi'(nd_i)=[(D^{r-1}\cdot C_i)/(r-1)!]n^{r-1}+\cdots\\ &h^1(nd_i)=0\end{aligned}$$

for $i=1,2$ and for n large. [The first formula, which gives the (well-known) leading term of the Hilbert polynomial of the ample divisor d_i, is a by-product of the proof; the second formula is one of Serre's classic results.] Then if n is large

$$\begin{aligned}\chi'(nD_0)-\chi'((n-1)D_0) &= h^0(nd_1)-h^0((n-1)d_2)\\ &= [(D^{r-1}\cdot C_1)n^{r-1}-(D^{r-1}\cdot C_2)(n-1)^{r-1}]/(r-1)!+\cdots\\ &= [(D^r\cdot V)/(r-1)!]n^{r-1}+\cdots.\end{aligned}$$

Therefore for n large $\chi'(nD_0)=[(D^r\cdot V)/r!]n^r+\cdots$. Again, since $(D^r\cdot V)>0$, $|nD_0|$ is not empty for large n.

Fix n large, and replace D_0 by a fixed member of $|nD_0|$, which we shall again denote by D_0. Let $\Delta_1,\cdots,\Delta_m$ be the irreducible components of the support of D_0, and let $\mathfrak{J}_1,\cdots,\mathfrak{J}_m$ be their defining sheaves of prime ideals. Fix q, $1\leqq q\leqq m$, and consider the inclusion $0\to\mathcal{O}(-D_0)\to\mathfrak{J}_q$. Its cokernel $\mathcal{K}$ has support in $\Delta_1\cup\cdots\cup\Delta_m$. So each of the minimal primes of the annihilator of $\mathcal{K}$ contains some $\mathfrak{J}_i$. Thus $\mathfrak{J}_1{}^k\cdots\mathfrak{J}_m{}^k\cdot\mathcal{K}=0$ for some integer k. Let $\mathfrak{J}_{ij}=\mathfrak{J}_1{}^{i+1}\cdots\mathfrak{J}_{j-1}{}^{i+1}\cdot\mathfrak{J}_j{}^i\cdots\mathfrak{J}_m{}^i$, $i=0,\cdots,k$, $j=1,\cdots,m$. The sheaves $\mathfrak{J}_{ij}\cdot\mathcal{K}$ ordered lexicographically filter $\mathcal{K}$. Their successive quotients $\mathfrak{Q}_{ij}$ are $\mathcal{O}_{\Delta_l}$-modules for some l, $1\leqq l\leqq m$, depending on i and j. So $\mathfrak{Q}_{ij}\otimes\mathcal{O}(pD_0)=\mathfrak{Q}_{ij}\otimes\mathcal{O}_{\Delta_l}\otimes\mathcal{O}(pD_0)=\mathfrak{Q}_{ij}\otimes\mathcal{O}(p\delta_l)$ where $\delta_l=D\cdot\Delta_l$. Then by induction we may assume $h^1(\mathfrak{Q}_{ij}\otimes\mathcal{O}(pD_0))=h^1(\mathfrak{Q}_{ij}\otimes\mathcal{O}(p\delta_l))=0$ for large p and all i and j. Hence $h^1(\mathcal{K}\otimes\mathcal{O}(pD_0))=0$. So the exact sequence

$$0\to\mathcal{O}((p-1)D_0)\to\mathfrak{J}_q\otimes\mathcal{O}(pD_0)\to\mathcal{K}\otimes\mathcal{O}(pD_0)\to 0$$

obtained from the exact sequence $0\to\mathcal{O}(-D_0)\to\mathfrak{J}_q\to\mathcal{K}\to 0$ by tensoring it with $\mathcal{O}(pD_0)$ yields the surjection

$$H^1(\mathcal{O}((p-1)D_0))\to H^1(\mathfrak{J}_q\otimes\mathcal{O}(pD_0))\to 0$$

for large p. Similarly the exact sequence

$$0\to\mathfrak{J}_q\otimes\mathcal{O}(pD_0)\to\mathcal{O}(pD_0)\to\mathcal{O}(p\delta_q)\to 0 \tag{1}$$

obtained from the exact sequence $0\to\mathfrak{J}_q\to\mathcal{O}_V\to\mathcal{O}_{\Delta_q}\to 0$ by tensoring it with $\mathcal{O}(pD_0)$ yields the surjection

$$H^1(\mathfrak{J}_q\otimes\mathcal{O}(pD_0))\to H^1(\mathcal{O}(pD_0))\to 0$$

for large p. Together, these two surjections imply that $h^1(\mathfrak{J}_q\otimes\mathcal{O}(pD_0))=h^1(\mathcal{O}(pD_0))=h$, a constant independent of p, for all large p. Therefore the long exact cohomology sequence associated with (1) shows that the trace

of $|pD_0|$ on Δ_q is complete. If V is a curve, it follows that $|pD_0|$ has no base points on Δ_q, $q=1,\cdots,m$, and therefore no base points at all for large p. If V is r-dimensional, $r>1$, then since we assume by induction that $|p\delta_q|$ has no base points, $q=1,\cdots,m$, also $|pD_0|$ has no base points for large p.

Fix n large, and let $T_n\colon V\to V_n$ be the regular rational transformation defined by the complete linear system $|nD_0|$. Let Q be any point of V_n, and set $Y=T_n^{-1}(Q)$. Let Δ be a general member of $|nD_0|$. Δ corresponds to a general hyperplane section C of V_n. Since $Q\notin C$, $\Delta\cap Y$ is empty. But if Y were of positive dimension i, then $(D^i\cdot Y)>0$, by assumption. Therefore, Y must be a finite set of points.

Finally we prove T_n is biregular at any point P of V for all large n. Then T_n will be biregular in a neighborhood of P and hence by quasi-compactness biregular on all of V for n large. Let $Q=T_n(P)$ and $Y=T_n^{-1}(Q)$. Since Y is a finite set of points and V is projective, Y is contained in an affine open subset $\mathfrak{U}$ of V. The complement X of $\mathfrak{U}$ is a closed subset of V. Hence its transform Z on V_n is also a (proper) closed subset. Z is contained in some hypersurface section S of high degree which misses Q. Let $\Delta\in|pD_0|$ correspond to S. Its complement $\mathfrak{V}$ contains P and is contained in the affine open set $\mathfrak{U}$. The ring $\mathcal{O}_P$ is a localization of the coordinate ring A of $\mathfrak{U}$, and A is a finitely generated algebra. Hence to complete the proof we note that for every $x\in A$, $(x)+q\Delta>0$ for some q depending on x. In fact, the question being local, we may work on some affine open set $\mathfrak{W}$ meeting Δ on which Δ is given by a single local equation g. But then, since x is defined on the complement of the set of zeros of g, xg^q is defined on all of W for some q.

2. Generalization to Theorem 2. First of all, we may assume V is a variety (i.e., an integral algebraic scheme). In fact, let V_i, $i=1,\cdots,m$, be the irreducible components of V given their reduced structure, and let $\alpha_i\colon V_i\to V$ be the corresponding inclusion morphisms. The natural morphism from the disjoint union of the V_i into V is finite and surjective. Hence by [1. III, 2.6.2] D is ample for V if all the $\alpha_i^*(D)$ are ample for V_i. But clearly if $\alpha\colon Z\to V$ is the inclusion of any closed subscheme Z of V, $\alpha^*(D)$ is arithmetically positive on Z. Hence if Theorem 2 holds for the V_i, which are integral, it holds for V.

We proceed to prove that $|nD|$ has no base points for large n. Again, we first prove that $|nD|$ is not empty for large n.

This time we cannot write $D=C_1-C_2$ where C_1 and C_2 are prime

divisors. Instead let $\mathfrak{I}_1 = \mathcal{O}(-D) \cap \mathcal{O}_V$ and $\mathfrak{I}_2 = \mathfrak{I}_1 \cdot \mathcal{O}(D)$. Then $\mathfrak{I}_1$ and $\mathfrak{I}_2$ are coherent sheaves of ideals. Let C_1 and C_2 be the closed subschemes they define, and let $\alpha_1 \colon C_1 \to V$ and $\alpha_2 \colon C_2 \to V$ be the corresponding inclusions.

Let $d_1 = \alpha_1{}^*(D)$ and $d_2 = \alpha_2{}^*(D)$. We may assume that $h^q(nd_1) = 0$ and $h^q(nd_2) = 0$ for $q > 0$ and n large by a dimension argument if V is a curve and by induction if V is higher dimensional. Then the exact sequences

$$\begin{array}{ccccccccc} 0 \to & \mathfrak{I}_1 \otimes \mathcal{O}(nD) & \to & \mathcal{O}(nD) & \to & \mathcal{O}(nd_1) & \to 0 \\ & \| & & & & & \\ 0 \to & \mathfrak{I}_2 \otimes \mathcal{O}((n-1)D) & \to & \mathcal{O}((n-1)D) & \to & \mathcal{O}((n-1)d_2) & \to 0 \end{array}$$

yield

$$\chi(nD) - \chi((n-1)D) = h^0(nd_1) - h^0((n-1)d_2)$$

where $\chi(nD) = h^0(nD) - h^1(nD) + h^2(nD) - \cdots$ and

$$\chi'(nD) - \chi'((n-1)D) = h^0(nd_1) - h^0((n-1)d_2)$$

where $\chi'(nD) = h^0(nD) - h^1(nD)$. So $\chi(nD) - \chi'(nD)$ is constant.[1] Since $\chi(nD) \to \infty$, also $\chi'(nD) \to \infty$. Therefore $|nD|$ is not empty if n is large.

We can finish the proof that $|nD|$ has no base points for large n exactly as above, or we may clean up part of that proof by applying the induction hypothesis directly to the subscheme defined by $\mathcal{K}$ to show that $h^1(\mathcal{K} \otimes \mathcal{O}(nD)) = 0$.

We show that the morphism $T_n \colon V \to V_n$ (n large) defined by $|nD|$ has finite fibres much as we did before. Let $Q \in V_n$ be any closed point, and set $Y = T_n{}^{-1}(Q)$. Then some hyperplane section C of V_n misses Q. C corresponds to a member Δ of $|nD|$. Since $Q \notin C$, $\Delta \cap Y$ is empty. Thus $\mathcal{O}(nD) \otimes \mathcal{O}_Y \cong \mathcal{O}_Y$, and so $\chi(\mathcal{O}(pD) \otimes \mathcal{O}_Y)$ is independent of p. Hence Y must be a finite set of points.

The birational correspondence between the normalizations W of V and W_n of V_n in $k(V)$ is finite-valued and hence an isomorphism. Therefore $\pi^*(D)$ is ample where $\pi \colon W \to V$ is the natural morphism. Since π is finite and surjective, D is ample for V by [1. III, 2.6.2].

Alternately, we may obtain an affine open subset of V containing Y by Chevalley's theorem [1. II, 6.7.1] and finish up as before. In fact, the

---

[1] The key difference between our proof and Nakai's appears here. In [5], Nakai ingeniously shows that if V is projective then $h^q(nD) = 0$ for $q > 1$ and n large. We show, in essence, that $h^q(nD)$ is constant in general (V being complete) for $q > 1$ and n large.

inverse image on W of any affine open $\mathfrak{U}$ on V_n is affine open, and therefore the inverse image on V of $\mathfrak{U}$ is also affine open.

HARVARD UNIVERSITY.

REFERENCES.

[1] A. Grothendieck, "Éléments de Géométrie Algébrique, I, II, III," *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, No. 4, 8, 11 (1960-1961).

[2] B. G. Moisezon, "Projective imbeddings of algebraic varieties," *Soviet Mathematics-Doklady*, vol. 2, No. 6 (Nov. 1961), pp. 1146-1148.

[3] ———, "Remarks on projective imbeddings of algebraic varieties," *Soviet Mathematics-Doklady*, vol. 3, No. 4 (July 1962), pp. 1500-1503.

[4] Y. Nakai, "Non-degenerate divisors on an algebraic surface," *Journal of Science of the Hiroshima University*, vol. 24 (1960).

[5] ———, "A criterion of an ample sheaf on a projective scheme," *American Journal of Mathematics*, vol. 85 (1963), pp. 14-26.

ON CHARACTERISTIC SYSTEMS OF FAMILIES OF SURFACES WITH ORDINARY SINGULARITIES IN A PROJECTIVE SPACE.*

By K. KODAIRA.

In our previous paper [5] we have proved a theorem of completeness of characteristic systems for analytic families of surfaces with ordinary singularities in ambient threefolds. In this paper we examine the application of the theorem to surfaces with ordinary singularities in a projective 3-space. The theorem asserts that the characteristic systems of a maximal analytic family of surfaces with ordinary singularities in an ambient threefold are complete if the surfaces are semi-regular. We show, in Section 2, that, in a projective 3-space, the semi-regularity coincides with the regularity and that the regularity is equivalent to the linear independence of certain simultaneous linear equations (Theorem 2). We reformulate the theorem of completeness of characteristic systems for analytic families of surfaces with ordinary singularities in a projective 3-space (Theorem 1). In Section 4 we prove some criteria of regularity (Theorems 5, 6 and 7). The problem of proving the completeness of the characteristic systems of complete continuous systems of surfaces with ordinary singularities in a projective 3-space has been proposed by O. Zariski (see Zariski [9], p. 99). Our results in Sections 2 and 4 may be considered as partial answers to this problem. In Section 5 we consider several concrete examples of surfaces. First we show with the aid of the criteria that certain classical examples of surfaces are regular. We then examine surfaces whose double curves are non-singular complete intersections and find that they are not regular except some special cases whereas they form maximal families with complete characteristic systems. Apparently this result indicates that the requirement of semi-regularity imposes a strong restriction on surfaces with ordinary singularities. Finally we examine the application of Theorem 1 to canonical surfaces of irregularity 0, of genus 4 and of order 7 studied earlier by Enriques [2] and by Maxwell [7].

1. Preliminaries. We denote by W a projective 3-space defined over the field $\boldsymbol{C}$ of complex numbers. We denote a point in W by $w = (w_0, w_1, w_2, w_3)$,

---

Received June 4, 1964.

* This work was partially supported by the National Science Foundation under Grant NSF-G7030.

where w_0, w_1, w_2, w_3 are the homogeneous coordinates of w in a fixed coordinate system. We cover W by a finite number of small polycylindrical neighborhoods W_i, $i \in I$, where I denotes a finite set of indices. We choose for each index i a coordinate $w_{\alpha(i)}$ which does not vanish on W_i and let

$$e_{ik} = e_{ik}(w) = w_{\alpha(k)}/w_{\alpha(i)} \qquad \text{for } w \in W_i \cap W_k.$$

Moreover we define E to be the complex line bundle over W determined by the system $\{e_{ik}\}$ of transition functions e_{ik}. It is well known that the group of complex line bundles over W is an infinite cyclic group generated by E. We denote by L_m the linear space consisting of all homogeneous polynomials $\phi = \phi(w)$ of degree m in w_0, w_1, w_2, w_3. We define

$$\phi_i = \phi_i(w) = \phi(w/w_{\alpha(i)}), \qquad \text{for } \phi \in L_m, \tag{1}$$

where $w/w_{\alpha(i)}$ stands for $(w_0/w_{\alpha(i)}, w_1/w_{\alpha(i)}, w_2/w_{\alpha(i)}, w_3/w_{\alpha(i)})$. Obviously ϕ_i is a holomorphic function on W_i and

$$\phi_i = e_{ik}{}^m \cdot \phi_k, \qquad \text{on } W_i \cap W_k.$$

Thus the collection $\{\phi_i\}$ of the functions ϕ_i, $i \in I$, represents a holomorphic section of the complex line bundle mE over W.

Let S be an algebraic surface of *order* n in W with *ordinary sigularities only* defined by an irreducible equation

$$f^0(w) = 0, \qquad f^0 \in L_n,$$

and let Δ be the *double curve* of S. Moreover let I_Δ denote the subset of I consisting of those indices i for which the W_i meet Δ. We choose on each neighborhood W_i, $i \in I_\Delta$, a local coordinate (x_i, y_i, z_i) such that $f^0{}_i(w) = f^0(w/w_{\alpha(i)})$ coincides with one of the polynomials $y_i z_i$, $x_i y_i z_i$, $x_i y_i{}^2 - 4 z_i{}^2$. This is possible provided that the neighborhoods W_i are sufficiently small. Thus we have three divisions I_d, I_t, I_c of the set I_Δ and

$$f^0{}_i(w) = \begin{cases} y_i z_i, & \text{for } i \in I_d, \\ x_i y_i z_i, & \text{for } i \in I_t, \\ x_i y_i{}^2 - 4 z_i{}^2, & \text{for } i \in I_c. \end{cases} \tag{2}$$

For $i \in I_c$ we denote by c_i the center: $x_i = y_i = z_i = 0$ of the polycylinder W_i. Note that the points c_i, $i \in I_c$, are the *cuspidal points* of the surface S.

In what follows we denote by $t = (t_1, t_2, \cdots, t_\mu)$ a point in the space C^μ of several complex variables and define

$$|t| = \max_\nu |t_\nu|.$$

Moreover we denote by M_ϵ the polycylinder in C^μ consisting of all points t with $|t| < \epsilon$. Consider an *analytic family* $\mathfrak{F}$ of algebraic surfaces S_t, $t \in M_1$, with ordinary singularities in W defined respectively by irreducible equations

$$f(w,t) = 0$$

such that S_0 coincides with S (see [5], § 2). Obviously we may assume that the coefficients of the homogeneous polynomial $f(w,t)$ in w are holomorphic functions of t and that

$$f(w,0) = f^0(w).$$

Let Δ_t denote the double curve of the surface S_t. *We are concerned with the restriction* $\mathfrak{F} \mid M_\epsilon$ *of* $\mathfrak{F}$ *to a small polycylinder* M_ϵ. Hence we may assume that Δ_t is covered by the neighborhoods W_i, $i \in I_\Delta$, and that, on each neighborhood W_i, $i \in I_\Delta$, there exist three independent holomorphic functions

$$X_i(t) = X_i(w,t), \quad Y_i(t) = Y_i(w,t), \quad Z_i(t) = Z_i(w,t)$$

depending holomorphically on t, $t \in M_1$, and satisfying the boundary conditions

$$X_i(w,0) = x_i, \quad Y_i(w,0) = y_i, \quad Z_i(w,0) = z_i,$$

such that

$$f_i(w,t) = \begin{cases} Y_i(t)Z_i(t), & \text{for } i \in I_d, \\ X_i(t)Y_i(t)Z_i(t), & \text{for } i \in I_t, \\ X_i(t)Y_i(t)^2 - 4Z_i(t)^2, & \text{for } i \in I_c, \end{cases} \tag{3}$$

where $f_i(w,t) = f_i(w/w_{\alpha(i)}, t)$. For $i \in I_c$ we denote by $\mathfrak{c}_{ti}$ the cuspidal point of the surface S_t determined by the equations $X_i(w,t) = Y_i(w,t) = Z_i(w,t) = 0$.

We define

$$L_m \mid S_t = L_m/L_{m-n}f(t), \qquad f(t) = f(w,t),$$

where $L_{m-n}f(t)$ denotes the linear subspace of L_m consisting of all multiples $gf(t)$, $g \in L_{m-n}$, of $f(t)$. We call $L_m \mid S_t$ the *restriction* of L_m to S_t. Moreover we denote by r_t the canonical homomorphism of L_m onto $L_m \mid S_t$ and call r_t the *restriction map* to S_t.

Let $L_n(-\Delta_t - \sum \mathfrak{c}_t')$ denote the linear subspace of L_n consisting of those homogeneous polynomials ϕ of degree n in w which vanish on Δ_t and satisfy

$$(\partial\phi_i/\partial Y_i(t))(\mathfrak{c}_{ti}) = 0, \qquad \text{for } i \in I_c, \tag{4}$$

where the left side of (4) denotes the value at $\mathfrak{c}_{ti}$ of the partial derivative

$\partial\phi_i/\partial Y_i(t)$ of ϕ_i *considered as a function of the local coordinates* $X_i(t)$, $Y_i(t)$, $Z_i(t)$ (compare [5], Definition 2). We define

$$\Lambda(S_t) = r_t L_n(-\Delta_t - \sum c_t').$$

Let $\partial/\partial t$ denote an arbitrary tangent vector of M_1 at t. We infer from (3) that the partial derivative $\partial f(t)/\partial t$ of $f(t) = f(w, t)$ belong to the linear space $L_n(-\Delta_t - \sum c_t')$. Hence $r_t[\partial f(t)/\partial t]$ is an element of the linear space $\Lambda(S_t)$. We call $r_t[\partial f(t)/\partial t]$ the infinitesimal displacement of S_t along the tangent vector $\partial/\partial t$ and denote it by the *symbol* $\partial S_t/\partial t$:

$$\partial S_t/\partial t = r_t[\partial f(t)/\partial t].$$

Moreover we denote by σ_t the *linear map* $\partial/\partial t \to \partial S_t/\partial t$ of the tangent space $T_t(M_1)$ of M_1 at t into the linear space $\Lambda(S_t)$ (compare [5], Definition 6). The range $\sigma_t T_t(M_1)$ of σ_t is called the *characteristic system of* $\mathfrak{F}$ *on* S_t. *The characteristic system* $\sigma_t T_t(M_1)$ *is said to be complete if and only if* $\sigma_t T_t(M_1) = \Lambda(S_t)$. We shall say that the family $\mathfrak{F}$ is *effectively parametrized* if and only if the map σ_t is injective for each point $t \in M_1$.

By a *maximal* analytic family of surfaces with ordinary singularities in W we shall mean an analytic family $\mathfrak{F}$ of algebraic surfaces S_t, $t \in M$, with ordinary singularities in W which is maximal at each point t of M, where M denotes a complex manifold. We recall that $\mathfrak{F}$ is said to be maximal at t if and only if, for any analytic family $\mathfrak{F}'$ of algebraic surfaces S_u', $u \in M'$, with ordinary singularities in W such that $S_o' = S_t$, there exist a neighborhood N' of o in M' and a holomorphic map h of N' into M with $h(o) = t$ such that $S_u' = S_{h(u)}$ for $u \in N'$ (see [5], Definition 5).

2. A theorem of completeness of characteristic systems. Let S be an algebraic surface of order n in W with ordinary singularities only defined by an irreducible equation $f^0(w) = 0$ and let Δ denote the double curve of S. Moreover let W_i, x_i, y_i, z_i, I_Δ, I_d, I_t, I_c, $\cdots$ have the same meaning as in § 1. We denote the number of the cuspidal points of the surface S by γ and indicate the cuspidal points by $c_1, c_2, \cdots, c_i, \cdots, c_\gamma$. Thus we suppose that I_c consists of the integers $1, 2, \cdots, \gamma$. Let $L_n(-\Delta)$ denote the linear space of those homogeneous polynomials of degree n in w which vanish on Δ. $L_n(-\Delta - \sum c')$ is, by definition, the linear subspace of $L_n(-\Delta)$ consisting of those homogeneous polynomials $\phi \in L_n(-\Delta)$ which satisfy the simultaneous linear equations $(\partial\phi_i/\partial y_i)(c_i) = 0$, $i = 1, 2, \cdots, \gamma$.

THEOREM 1 (*completeness of characteristic systems*). *If the linear equations*

$$(\partial\phi_i/\partial y_i)(c_i) = 0, \qquad i = 1, 2, \cdots, \gamma,$$

are linearly independent on the linear space $L_n(-\Delta)$, *then there exists a maximal analytic family* $\mathfrak{F}$ *of algebraic surfaces* S_t, $t \in M_1$, *with ordinary singularities in* W *such that* S_0 *coincides with* S *and such that, for each point* $t \in M_1$, *the map* $\sigma_t\colon \partial/\partial t \to \partial S_t/\partial t$ *maps the tangent space* $T_t(M_1)$ *of* M_1 *isomorphically onto* $\Delta(S_t)$. *Thus the family* $\mathfrak{F}$ *is effectively parametrized and the characteristic system of* $\mathfrak{F}$ *on each surface* S_t, $t \in M_1$, *is complete.*

We shall derive this theorem from a result of [5]. Consider a complex line bundle mE over W, where m is an arbitrary integer. For any holomorphic section $\phi\colon w \to \phi(w)$ of mE defined on a domain $\mathfrak{D} \subseteq W$, we denote by $\phi_i(w)$ the *fibre coordinate* of $\phi(w)$ over W_i. Note that $\phi_i(w)$ is a holomorphic function of w defined on $\mathfrak{D} \cap W_i$ and that

$$\phi_i(w) = (e_{ik}(w))^m \phi_k(w), \qquad \text{on } \mathfrak{D} \cap W_i \cap W_k.$$

Let $\Omega(mE)$ denote the sheaf over W of germs of holomorphic sections of mE and let $\Omega(mE-\Delta)$ be the subsheaf of $\Omega(mE)$ consisting of germs of those holomorphic sections which vanish on Δ. Moreover let $\Omega(mE-\Delta-\sum \mathfrak{c}')$ denote the subsheaf of $\Omega(mE-\Delta)$ consisting of germs of those holomorphic sections ϕ of mE which vanish on Δ and satisfy

$$(\partial\phi_i/\partial y_i)(\mathfrak{c}_i) = 0, \qquad i = 1, 2, \cdots, \gamma.$$

We set

$$\Phi = \Omega(nE - \Delta - \textstyle\sum \mathfrak{c}'),$$

where n denote the order of the surface S.

Definition. We say that the surface S is regular in W if and only if the cohomology group $H^1(W, \Phi)$ vanishes (compare Zariski [9], p. 67).

We define

$$\Psi = r\Phi$$

where r denotes the restriction map to the surface S (compare [5], p. 599). The map r induces a homomorphism

$$r^*\colon H^1(W, \Phi) \to H^1(S, \Psi).$$

The surface S is said to be *semi-regular* in W if and only if $r^* H^1(W, \Phi)$ vanishes (see [5], Definition 7). We have the exact sequence

$$0 \to \Omega \to \Phi \xrightarrow{r} \Psi \to 0, \tag{5}$$

where Ω denotes the sheaf over W of germs of holomorphic functions. Since

$H^{\nu}(W,\Omega)$ vanishes for $\nu=1,2,3,\cdots$, we infer from (5) that $r^*: H^1(W,\Phi) \to H^1(S,\Psi)$ is bijective. Hence the surface S is semi-regular in W if and only if $H^1(W,\Phi)$ vanishes. Thus, *in the projective space W, the semi-regularity coincides with the regularity.* We obtain from (5) the equality

$$\dim \Delta(S) = \dim L_n(-\Delta-\sum \mathfrak{c}') - 1. \tag{6}$$

We construct a non-singular model $\tilde{S}$ of S in an *obvious manner* and denote the natural holomorphic map of $\tilde{S}$ onto S by h. The non-singular model $\tilde{S}$ is characterized by the following properties: i) Let q be a point on $\tilde{S}$. If the point $p=h(q)$ on S is not cuspidal, then h maps a neighborhood U of q on $\tilde{S}$ *biholomorphically into* W. If $p=h(q)$ is a cuspidal point of S, then there exist a local coordinate (u,v) on $\tilde{S}$ with the center q and a local coordinate (x,y,z) on W with the center p such that, in a neighborhood of q, the map h is written in the form

$$h:(u,v)\to(x,y,z)=(u^2,v,\tfrac{1}{2}uv).$$

Let $\tilde{E}$ be the complex line bundle over $\tilde{S}$ induced from the complex line bundle E over W by the map h and let $\Omega(m\tilde{E})$ denote the sheaf over $\tilde{S}$ of germs of holomorphic sections of $m\tilde{E}$. Moreover let $\tilde{\Delta}=h^{-1}(\Delta)$ and denote by $\Omega(m\tilde{E}-\tilde{\Delta})$ the subsheaf of $\Omega(m\tilde{E})$ consisting of germs of those holomorphic sections of $m\tilde{E}$ which vanish on $\tilde{\Delta}$. Consider the restriction $r\Omega(mE-\Delta)$ of $\Omega(mE-\Delta)$ to the surface S. We infer readily that *the map h induces an isomorphism of the stalk $r\Omega_p(mE-\Delta)$ of $r\Omega(mE-\Delta)$ over a point $p\in S$ onto the direct sum of the stalks $\Omega_q(m\tilde{E}-\tilde{\Delta})$, $q\in h^{-1}(p)$, of $\Omega(m\tilde{E}-\tilde{\Delta})$* (see [5], §1). Thus we obtain the *isomorphism*

$$H^{\nu}(S,r\Omega(mE-\Delta))\cong H^{\nu}(\tilde{S},\Omega(m\tilde{E}-\tilde{\Delta})), \qquad \nu=0,1,2. \tag{7}$$

Now we show that the characteristic class $c(\tilde{E})$ of the complex line bundle $\tilde{E}$ is *positive* in the sense that there exists on $\tilde{S}$ a Kähler form $\tilde{\omega}$ belonging to $c(\tilde{E})$. It is well known that the standard Kähler form ω on W belongs to the characteristic class $c(E)$ of E. It follows that the closed form $h^*\omega$ of type $(1,1)$ induced on $\tilde{S}$ belongs to the characteristic class $c(\tilde{E})$ of $\tilde{E}$. Clearly $h^*\omega$ is positive definite outside the points $h^{-1}(\mathfrak{c}_i)$, $i=1,2,\cdots,\gamma$, while we can find a real-valued C^∞-differentiable function g on $\tilde{S}$ such that $(-1)^{\frac{1}{2}}\partial\bar{\partial}g$ is positive definite at each point $h^{-1}(\mathfrak{c}_i)$, $i=1,2,\cdots,\gamma$. Hence, setting

$$\tilde{\omega}=h^*\omega+\epsilon(-1)^{\frac{1}{2}}\partial\bar{\partial}g, \qquad \epsilon>0,$$

we obtain a Kähler form $\tilde{\omega}$ belonging to $c(\tilde{E})$, provided that ϵ is sufficiently small.

For any divisor D on $\bar{S}$ we denote by $[D]$ the complex line bundle over $\bar{S}$ determined by D. By the canonical bundle $\bar{K}$ over $\bar{S}$ we shall mean the complex line bundle determined by a canonical divisor on $\bar{S}$. We have the classical formula:

$$\bar{K} = (n-4)\bar{E} - [\bar{\Delta}]. \tag{8}$$

THEOREM 2. *The surface S is regular in W if and only if the linear equations*

$$(\partial\phi_i/\partial y_i)(\mathfrak{c}_i) = 0, \qquad i = 1, 2, \cdots, \gamma, \tag{9}$$

are linearly independent on the linear space $L_n(-\Delta)$.

Proof. First we show that

$$H^{\nu}(W, \Omega(nE - \Delta)) = 0, \qquad \text{for } \nu = 1, 2. \tag{10}$$

From the exact sequence

$$0 \to \Omega \to \Omega(nE - \Delta) \to r\Omega(nE - \Delta) \to 0$$

we obtain the isomorphism

$$H^{\nu}(W, \Omega(nE - \Delta)) \cong H^{\nu}(S, r\Omega(mE - \Delta)), \qquad \text{for } \nu = 1, 2.$$

Combining this with (7) and (8), we get

$$H^{\nu}(W, \Omega(nE - \Delta)) \cong H^{\nu}(\bar{S}, \Omega(\bar{K} + 4\bar{E})), \qquad \text{for } \nu = 1, 2,$$

while, as was shown above, the characteristic class $c(\bar{E})$ is positive and therefore the cohomology group $H^{\nu}(\bar{S}, \Omega(\bar{K} + 4\bar{E}))$ vanishes for $\nu = 1, 2$ (see [3], Theorem 3). Consequently we obtain (10).

Setting $Q = \Omega(nE - \Delta)/\Phi$ we obtain the exact sequence

$$0 \to \Phi \to \Omega(nE - \Delta) \to Q \to 0. \tag{11}$$

Note that the stalk Q_w over a point w is given by the formula

$$Q_w = \begin{cases} \boldsymbol{C}, & \text{if } w = \mathfrak{c}_i, \\ 0, & \text{otherwise.} \end{cases}$$

Combining the exact cohomology sequence derived from (11) with (10) we obtain the exact sequence

$$0 \to L_n(-\Delta - \textstyle\sum \mathfrak{c}') \to L_n(-\Delta) \to \boldsymbol{C}^{\gamma} \to H^1(W, \Phi) \to 0, \tag{12}$$

where $\boldsymbol{C}^{\gamma}$ denotes the space of γ complex variables. (12) shows that S is

regular in W if and only if the linear equations (9) are linearly independent on $L_n(-\Delta)$, q. e. d.

Now we derive Theorem 1 from a result of [5] which is formulated as follows:

THEOREM 1*. *If the surface S is semi-regular in W, then there exists an analytic family $\mathfrak{F}$ of surfaces S_t, $t \in M_1$, with ordinary singularities in W such that S_0 coincides with S and such that, for each point $t \in M_1$, the map $\sigma_t: \partial/\partial t \to \partial S_t/\partial t$ maps $T_1(M)$ isomorphically onto $\Lambda(S_t)$. The family $\mathfrak{F}$ is maximal at $t=0$* (see [5], § 3).

Assume that the linear equations (9) are linearly independent on $L_n(-\Delta)$. Then, by Theorem 2, the surface S is regular and therefore semi-regular in W. Hence, by Theorem 1*, there exists an analytic family $\mathfrak{F}$ of surfaces S_t, $t \in M_1$, with ordinary singularities in W such that $S_0 = S$ and such that σ_t maps $T_t(M_1)$ isomorphically onto $\Lambda(S_t)$. For our purpose it suffices to show that the family $\mathfrak{F}$ is maximal at each point t, $|t| < \epsilon$, where ϵ is a small positive number.

First we calculate the dimension of the linear space $L_n(-\Delta)$. From the exact sequence

$$0 \to \Omega \to \Omega(nE - \Delta) \to r\Omega(nE - \Delta) \to 0$$

we get the exact sequence

$$0 \to \boldsymbol{C} \to L_n(-\Delta) \to H^0(S, r\Omega(nE - \Delta)) \to 0.$$

Combining this with (7) and (8) we obtain

$$\dim L_n(-\Delta) = 1 + \dim H^0(\tilde{S}, \Omega(\tilde{K} + 4\tilde{E})),$$

while the characteristic class $c(\tilde{E})$ is positive. Hence, using a theorem of Riemann-Roch (see [4], p. 123), we obtain

$$\dim L_n(-\Delta) = \pi(4\tilde{E}) + p_a + 1,$$

where $\pi(4\tilde{E})$ denotes the virtual genus of $4\tilde{E}$ and p_a is the arithmetic genus of $\tilde{S}$. Employing the notation of [4], we have

$$\pi(4\tilde{E}) = 8(\tilde{E}^2) + 2(\tilde{K}\tilde{E}) + 1.$$

Hence, denoting by d the order of Δ and using (8), we obtain

$$\dim L_n(-\Delta) = 2n^2 - 4d + p_a + 2. \tag{13}$$

Let $f(w,t), \Delta_t, c_{ti}, X_i(t), Y_i(t), \cdots$ have the same meaning as in the preceding section with respect to the family $\mathfrak{F}$ of the surfaces S_t. Since the non-singular models $\tilde{S}_t$, $t \in M_1$, form an analytic family, the arithmetic genus of $\tilde{S}_t$ is independent of t. Applying the formula (13) to the surface S_t, we infer therefore that *the dimension of* $L_n(-\Delta_t)$ *is independent of* t. Consider the linear equations

$$(l)_t \qquad (\partial\phi_i/\partial Y_i(t))(c_{ti}) = 0, \qquad i = 1, 2, \cdots, \gamma.$$

By hypothesis the linear equations $(l)_0$ are linearly independent on $L_n(-\Delta_0)$, while the dimension of $L_n(-\Delta_t)$ is independent of t. Hence, for each t with $|t| < \epsilon$, the linear equations $(l)_t$ are independent on $L_n(-\Delta_t)$, provided that the positive number ϵ is sufficiently small, and therefore, by Theorem 2, each surface S_t, $|t| < \epsilon$, is regular in W.

Take a point s, $|s| < \epsilon$, of M_1. Since the surface S_s is semi-regular in W, there exists, by Theorem 1*, an analytic family $\mathfrak{F}'$ of surfaces S'_u, $u \in M'$, with ordinary singularities in W such that $S'_0 = S_s$ and such that $\sigma' \colon \partial/\partial u \to [\partial S'_u/\partial u]_{u=0}$ maps $T_0(M')$ isomorphically onto $\Lambda(S'_0)$, where M' is a polycylinder in the space of several complex variables $u_1, \cdots, u_\nu, \cdots, u_\beta$ containing the origin 0. Moreover $\mathfrak{F}'$ is maximal at $u = 0$. Hence there exist a neighborhood N of s in M_1 and a holomorphic map: $t \to u = u(t)$ of N into M' with $u(s) = 0$ such that $S_t = S'_{u(t)}$. Let

$$f'(w, u) = 0$$

be the irreducible equation of S'_u, where $f'(w,u)$ denotes a homogeneous polynomial of degree n in w whose coefficients are holomorphic functions of u. The equality $S_t = S'_{u(t)}$ implies that

$$f(w,t) = c(t) f'(w, u(t)).$$

Hence we obtain

$$[\partial S_t/\partial t]_{t=s} = c(s) \sum_{\nu=1}^{\beta} [\partial u_\nu(t)/\partial t]_{t=s} [\partial S'_u/\partial u_\nu]_{u=0},$$

while the map $\sigma_s \colon \partial/\partial t \to [\partial S_t/\partial t]_{t=s}$ maps $T_s(M_1)$ isomorphically onto $\Lambda(S_s) = \Lambda(S'_0)$. We infer therefore that the map $t \to u(t)$ maps a neighborhood of s in M_1 biholomorphically onto a neighborhood of 0 in M'. It follows immediately that the family $\mathfrak{F}$ is maximal at s, q. e. d.

3. An exact sequence. In this section we shall derive an exact sequence of sheaves over a projective space. Let W_n denote a projective space of n dimensions and let Ξ_n be the sheaf over W_n of germs of holomorphic

vector fields. Moreover let W_{n-1} be a hyperplane in W_n and let Ξ_{n-1} be the sheaf over W_{n-1} of germs of holomorphic vector fields. Let E_n denote the complex line bundle over W_n determined by the divisor W_{n-1} and let E_{n-1} be the restriction of E_n to W_{n-1}. We write $E_n{}^2$ for the tensor product $E_n \otimes E_n$. Thus we consider in this section the group of complex line bundles over W_n as a multiplicative group. We indicate the Whitney sum of vector bundles by the symbol $\oplus$. We consider the sheaf

$$\Omega(E_n{}^2 \oplus E_n{}^2 \oplus \cdots \oplus E_n{}^2)$$

over W_n of germs of holomorphic sections of the Whitney sum of n copies of the line bundle $E_n{}^2$.

THEOREM 3. *We have the exact sequence*

$$0 \to \Xi_n \xrightarrow{\iota} \Omega(E_n{}^2 \oplus E_n{}^2 \oplus \cdots \oplus E_n{}^2) \xrightarrow{\kappa} E_{n-1} \otimes \Xi_{n-1} \to 0, \tag{14}$$

where $E_{n-1} \otimes \Xi_{n-1}$ stands for the tensor product $\Omega(E_{n-1}) \otimes_\Omega \Xi_{n-1}$.

Proof. We denote a point in W_n by $w = (w^0, w^1, \cdots, w^n)$, where $w^0, w^1, \cdots, w^n$ are the homogeneous coordinates of w in a fixed coordinate system. Let U_α be the open subset of W_n consisting of all points w with $w^\alpha \neq 0$ and define

$$z_\alpha{}^\lambda = w^\lambda / w^\alpha, \qquad \text{on } U_\alpha.$$

On the intersection $U_\alpha \cap U_\beta$, $\alpha \neq \beta$, we have

$$\begin{cases} z_\alpha{}^\beta = 1/z_\beta{}^\alpha, \\ z_\alpha{}^\lambda = z_\beta{}^\lambda / z_\beta{}^\alpha, & \text{for } \lambda \neq \alpha, \neq \beta. \end{cases} \tag{15}$$

The complex line bundle E_n is defined by the system $\{e_{\alpha\beta}\}$ of transition functions $e_{\alpha\beta} = w^\beta / w^\alpha$ with respect to the covering $\{U_\alpha\}$.

Consider a holomorphic vector field ξ on a domain $\mathfrak{D} \subseteq W_n$ and write ξ in the form

$$\xi = \sum_{\lambda \neq \alpha} \xi_\alpha{}^\lambda \partial / \partial z_\alpha{}^\lambda, \qquad \text{on } \mathfrak{D} \cap U_\alpha.$$

On the intersection $\mathfrak{D} \cap U_\alpha \cap U_\beta$ we have

$$\begin{cases} \xi_\alpha{}^\beta = -(z_\alpha{}^\beta)^2 \xi_\beta{}^\alpha, \\ \xi_\alpha{}^\lambda = (z_\alpha{}^\beta)^2 (z_\beta{}^\alpha \xi_\beta{}^\lambda - z_\beta{}^\lambda \xi_\beta{}^\alpha). \end{cases} \tag{16}$$

For each index $\lambda = 1, 2, \cdots, n$, we set

$$\begin{cases} \psi^\lambda{}_0 = \xi_0{}^\lambda, \\ \psi^\lambda{}_\lambda = -\xi_\lambda{}^0, \\ \psi^\lambda{}_\alpha = z_\alpha{}^0 \xi_\alpha{}^\lambda - z_\alpha{}^\lambda \xi_\alpha{}^0, & \text{for } \alpha \neq \lambda, \neq 0. \end{cases} \tag{17}$$

Using (16) we verify that

$$\psi^\lambda{}_\alpha = (e_{\alpha\beta})^2\psi^\lambda{}_\beta \qquad \text{on } \mathfrak{D} \cap U_\alpha \cap U_\beta.$$

Thus the collection $\psi^\lambda = \{\psi^\lambda{}_\alpha\}$ of $\psi^\lambda{}_\alpha$ represents a holomorphic section of $E_n{}^2$ over $\mathfrak{D}$. We set

$$\iota\xi = (\psi^1, \cdots, \psi^\lambda, \cdots, \psi^n).$$

Obviously the map $\xi \to \iota\xi$ defines an injection ι of Ξ into

$$\Omega(E_n{}^2 \oplus E_n{}^2 \oplus \cdots \oplus E_n{}^2).$$

Suppose that W_{n-1} is the hyperplane in W_n defined by the equation: $w^0 = 0$ and denote by r the restriction map to W_{n-1}. For any holomorphic section $\phi = (\phi^1, \cdots, \phi^\lambda, \cdots, \phi^n)$ of $E_n{}^2 \oplus \cdots \oplus E_n{}^2$ over a domain $\mathfrak{D} \subseteq W_n$ we define

$$\eta_\alpha{}^\lambda = r(\phi^\lambda{}_\alpha - z_\alpha{}^\lambda\phi^\alpha{}_\alpha), \qquad \lambda \neq \alpha, \alpha \neq 0,$$

where $\phi^\lambda{}_\alpha$ denotes the fibre coordinate of ϕ over U_α. We have

$$\phi^\lambda{}_\alpha = (e_{\alpha\beta})^2\phi^\lambda{}_\beta, \qquad \text{on } \mathfrak{D} \cap U_\alpha \cap U_\beta.$$

Hence, on the intersection $W_{n-1} \cap \mathfrak{D} \cap U_\alpha \cap U_\beta$, $\alpha \neq \beta$, we obtain

$$\begin{cases} \eta_\alpha{}^\beta = -e_{\alpha\beta}(z_\alpha{}^\beta)^2\eta_\beta{}^\alpha, \\ \eta_\alpha{}^\lambda = e_{\alpha\beta}(z_\alpha{}^\beta)^2(z_\beta{}^\alpha\eta_\beta{}^\lambda - z_\beta{}^\lambda\eta_\beta{}^\alpha), \quad \text{for } \lambda \neq \alpha, \neq \beta. \end{cases}$$

It follows that the collection of $\eta_\alpha{}^\lambda$ represents a section η of the sheaf $E_{n-1} \otimes \Xi_{n-1}$ over $W_{n-1} \cap \mathfrak{D}$ (compare (16)). We set

$$\kappa\phi = \eta.$$

Clearly the map $\phi \to \kappa\phi$ defines a homomorphism κ of $\Omega(E_n{}^2 \oplus \cdots \oplus E_n{}^2)$ onto $E_{n-1} \otimes \Xi_{n-1}$. The simultaneous equations

$$\begin{cases} \phi^\lambda{}_0 = \xi_0{}^\lambda, \\ \phi^\lambda{}_\lambda = -\xi_\lambda{}^0, \\ \phi^\lambda{}_\alpha = z_\alpha{}^0\xi_\alpha{}^\lambda - z_\alpha{}^\lambda\xi_\alpha{}^0, \end{cases} \qquad \alpha \neq \lambda, \neq 0,$$

are equivalent to

$$\begin{cases} \xi_0{}^\lambda = \phi^\lambda{}_0, \\ \xi_\alpha{}^0 = -\phi^\alpha{}_\alpha, \\ z_\alpha{}^0\xi_\alpha{}^\lambda = \phi^\lambda{}_\alpha - z_\alpha{}^\lambda\phi^\alpha{}_\alpha. \end{cases}$$

Hence we conclude the exactness of the sequence (14).

4. Some criteria of regularity. Let S be an algebraic surface of order n in W with ordinary singularities only and let Δ be the double curve of S. We denote by Ξ the sheaf over W of germs of holomorphic vector fields and by $\Xi(-\Delta)$ the subsheaf of Ξ consisting of germs of those holomorphic vector fields which vanish on Δ. Moreover we denote by Σ the subsheaf of Ξ consisting of germs of those holomorphic vector fields which are tangential to the curve Δ and vanish at the triple points of Δ. Obviously $\Xi(-\Delta)$ is a subsheaf of Σ. We define

$$r_\Delta\Xi = \Xi/\Xi(-\Delta),$$
$$\theta = \Sigma/\Xi(-\Delta),$$
$$N = \Xi/\Sigma.$$

Note that $r_\Delta\Xi$, θ and N can be considered as sheaves on the curve Δ. We call θ and N respectively the sheaves of *tangent vectors* and *normal vectors* on the curve Δ. We have the exact sequence

$$0 \to \theta \to r_\Delta\Xi \to N \to 0. \tag{18}$$

Let $\Omega(nE-2\Delta)$ denote the sheaf over W of germs of holomorphic sections of nE of which the fibre coordinates vanish on Δ together with their first partial derivatives. Clearly $\Omega(nE-2\Delta)$ is a subsheaf of the sheaf

$$\Phi = \Omega(nE-\Delta-\sum \mathfrak{c}').$$

THEOREM 4. *The sequence*

$$0 \to \Omega(nE-2\Delta) \to \Phi \to N \to 0 \tag{19}$$

is exact.

Proof. Let $f^0(w)=0$ be the irreducible equation of the surface S and let W_i, x_i, y_i, z_i, $\mathfrak{c}_i$, I_Δ, I_d, $\cdots$ have the same meaning as in Section 1 with respect to the surface S. By a *local* function and a *local* section on W [or W_i] we shall mean respectively a function and a section defined on a subdomain of W [or W_i]. We define

$$\Phi^\# = \Phi/\Omega(nE-2\Delta)$$

and for any local section ϕ of Φ we denote the corresponding local section of $\Phi^\#$ by $\phi^\#$. The fibre coordinate ϕ_i of ϕ on W_i is a local section of the sheaf $\Omega(-\Delta-\sum \mathfrak{c}')$ and

$$\phi_i = (e_{ik})^n \phi_k.$$

It follows that the fibre coordinate $\phi_i^{\#}$ of $\phi^{\#}$ on W_i is a local section of the quotient sheaf

$$\Omega(-\Delta-\sum c')/\Omega(2\Delta)$$

and

$$\phi_i^{\#}=(e_{ik})^n\phi_k^{\#}. \tag{20}$$

In what follows we denote by $\xi_i, \eta_i, \cdots$ local holomorphic functions on W_i and write r for the restriction map r_Δ to the curve Δ. First we consider the case in which Δ is non-singular. For $i\in I_d$ the fibre coordinate ϕ_i of any local section ϕ of Φ has the form

$$\phi_i=\zeta_i y_i+\eta_i z_i$$

and the first partial derivatives of ϕ_i vanish on Δ if and only if ζ_i and η_i both vanish on Δ. Hence we get

$$\phi_i^{\#}=r\zeta_i y_i^{\#}+r\eta_i z_i^{\#}. \tag{21}$$

For $i\in I_c$ we have

$$\phi_i=\xi_i y_i^2+2\eta_i x_i y_i-8\zeta_i z_i.$$

The first partial derivatives of ϕ_i vanish on Δ if and only if η_i and ζ_i vanish on Δ. Hence we obtain

$$\phi_i^{\#}=2r\eta_i x_i y_i^{\#}-8r\zeta_i z_i^{\#}. \tag{22}$$

Thus $\phi_i^{\#}$ is represented by a local holomorphic vector field $(r\eta_i, r\zeta_i)$ on the curve Δ. To prove that $\Phi^{\#}$ is isomorphic to N it suffices therefore to verify that

$$\begin{cases} r\eta_i\partial y_k/\partial y_i+r\zeta_i\partial y_k/\partial z_i=r\eta_k, \\ r\eta_i\partial z_k/\partial y_i+r\zeta_i\partial z_k/\partial z_i=r\zeta_k. \end{cases} \tag{23}$$

We have

$$f^0_i=\begin{cases} y_i z_i, & \text{for } i\in I_d, \\ x_i y_i^2-4z_i^2. & \text{for } i\in I_c. \end{cases}$$

It follows that

$$\phi_i^{\#}=(\eta_i\partial f^0_i/\partial y_i+\zeta_i\partial f^0_i/\partial z_i)^{\#}.$$

Substituting $(e_{ik})^n f^0_k$ for f^0_i we obtain

$$\phi_i^{\#}=(e_{ik})^n(\eta_i\partial f^0_k/\partial y_i+\zeta_i f^0_k/\partial z_i)^{\#},$$

while

$$\phi_i^{\#}=(e_{ik})^n\phi_k^{\#}=(e_{ik})^n(\eta_k\partial f^0_k/\partial y_k+\zeta_k\partial f^0_k/\partial z_k)^{\#}.$$

Hence we get

$$(\eta_i \partial f^0{}_k/\partial y_i + \zeta_i \partial f^0{}_k/\partial z_i)^\# = (\eta_k \partial f^0{}_k/\partial y_k + \zeta_k \partial f^0{}_k/\partial z_k)^\#.$$

Assuming that $k \in I_d$ we obtain therefore

$$\begin{aligned} r(\eta_i \partial y_k/\partial y_i + \zeta_i \partial y_k/\partial z_i) z_k{}^\# + r(\eta_i \partial z_k/\partial y_i \\ + \zeta_i \partial z_k/\partial z_i) y_k{}^\# = r\eta_k z_k{}^\# + r\zeta_k y_k{}^\#. \end{aligned}$$

This proves (23).

Now we consider the general case in which Δ has triple points. In case $i \in I_t$ the fibre coordinate ϕ_i of any local section ϕ of Φ has the form

$$\phi_i = \xi_i y_i z_i + \eta_i z_i x_i + \zeta_i x_i y_i.$$

Clearly the first partial derivative of ϕ_i vanish on Δ if and only if

$$\begin{aligned} \xi_1 &= \xi'_i x_i + \xi''_i y_i z_i, \\ \eta_i &= \eta'_i y_i + \eta''_i z_i x_i, \\ \zeta_i &= \zeta'_i z_i + \zeta''_1 x_i y_i, \end{aligned}$$

where $\xi'_i, \xi''_i, \eta'_i, \cdots$ are local holomorphic functions. In W_i the curve Δ is composed of three components Δ_x, Δ_y, Δ_z defined respectively by the equations $y_i = z_i = 0$, $z_i = x_i = 0$, $x_i = y_i = 0$. Denoting by r_x, r_y, r_z respectively the restriction maps to Δ_x, Δ_y, Δ_z, we obtain

$$\begin{aligned} \phi_i{}^\# = r_x\eta_i \cdot (x_i z_i)^\# + r_x\zeta_i \cdot (x_i y_i)^\# + r_y\zeta_i \cdot (x_i y_i)^\# \\ + r_y\xi_i \cdot (y_i z_i)^\# + r_z\xi_i \cdot (y_i z_i)^\# + r_z\eta_i \cdot (z_i x_i)^\#. \end{aligned} \tag{24}$$

Thus $\phi_i{}^\#$ is represented by the holomorphic vector fields $(r_x\eta_i, r_x\zeta_i)$, $(r_y\zeta_i, r_y\xi_i)$, $(r_z\xi_i, r_z\eta_i)$ defined respectively on the components Δ_x, Δ_y, Δ_z which satisfy the conditions

$$r_y\xi_i = r_z\xi_i, \quad r_z\eta_i = r_x\eta_i, \quad r_x\zeta_i = r_y\zeta_i, \qquad \text{at } x_i = y_i = z_i = 0.$$

Moreover if, for instance, $\Delta_x \cap W_i \cap W_k$ is non-empty, $k \in I_d$, we have

$$\phi_i{}^\# = (\eta_i \partial f^0{}_i/\partial y_i + \zeta_i \partial f^0{}_i/\partial z_i)^\#, \qquad \text{on } \Delta_x \cap W_i \cap W_k.$$

Comparing this with

$$\phi_k{}^\# = (\eta_k \partial f^0{}_k/\partial y_k + \zeta_k \partial f^0{}_k/\partial z_k)^\#,$$

we obtain

$$\begin{aligned} r_x\eta_i \partial y_k/\partial y_i + r_x\zeta_i \partial y_k/\partial z_i = r\eta_k, \\ r_x\eta_i \partial z_k/\partial y_i + r_x\zeta_i \partial z_k/\partial z_i = r\zeta_k. \end{aligned}$$

Thus we conclude that $\Phi^\#$ is isomorphic to N, q. e. d.

It is well-known that the cohomology groups $H^\nu(W, \Omega(kE - 2\Delta))$, $\nu = 1$,

$2,\cdots$, vanish for sufficiently large integers k. We define $n_0(\Delta)$ to be the minimum integer n_0 such that

$$H^{\nu}(W,\Omega(kE-2\Delta))=0, \qquad \text{for } k\geqq n_0, \nu=1,2. \tag{25}$$

THEOREM 5. *Assume that $n\geqq n_0(\Delta)$. The surface S is regular in W if and only if the cohomology group $H^1(\Delta,N)$ vanishes.*

Proof. In view of (25) we obtain from (19) the isomorphism

$$H^1(W,\Phi)\cong H^1(\Delta,N). \tag{26}$$

Hence the theorem follows.

We obtain from (18) the *exact sequence*

$$\cdots\to H^1(\Delta,\theta)\to H^1(\Delta,r_\Delta\Xi)\to H^1(\Delta,N)\to 0.$$

Hence *the vanishing of $H^1(\Delta,r_\Delta\Xi)$ implies the vanishing of $H^1(\Delta,N)$.* We have, by Theorem 3, the exact sequence

$$0\to\Xi\to\Omega(E^2\oplus E^2\oplus E^2)\xrightarrow{\kappa}E_2\otimes\Xi_2\to 0. \tag{27}$$

Obviously we may assume that the curve Δ meets the plane on which the sheaf Ξ_2 is defined transversally at d points $p_1,p_2,\cdots,p_d$, where d denotes the *order of* Δ. Now, restricting (27) to the curve Δ, we obtain the exact sequence

$$0\to r_\Delta\Xi\to r_\Delta\Omega(E^2\oplus E^2\oplus E^2)\xrightarrow{\kappa}\omega\to 0. \tag{28}$$

Note that the sheaf ω is defined over the set of the points $p_1,p_2,\cdots,p_d$ and each stalk of ω is isomorphic to $\boldsymbol{C}^2$. For any space $\mathscr{L}$ we denote by $\mathscr{L}^{\nu}$ the *direct sum of ν copies of $\mathscr{L}$.* Thus we rewrite (28) in the form

$$0\to r_\Delta\Xi\to[r_\Delta\Omega(2E)]^3\xrightarrow{\kappa}\omega\to 0. \tag{28$'$}$$

We obtain from this the exact sequence

$$\begin{aligned}0&\to H^0(\Delta,r_\Delta\Xi)\to[H^0(\Delta,r_\Delta\Omega(2E))]^3\xrightarrow{\kappa}\boldsymbol{C}^{2d}\\&\to H^1(\Delta,r_\Delta\Xi)\to[H^1(\Delta,r_\Delta\Omega(2E))]^3\to 0.\end{aligned} \tag{29}$$

It follows that $H^1(\Delta,r_\Delta\Xi)$ vanishes if the following two conditions are fulfilled:

α) *the cohomology group $H^1(\Delta,r_\Delta\Omega(2E))$ vanishes;*

β) *the map $\kappa:[H^0(\Delta,r_\Delta\Omega(2E))]^3\to\boldsymbol{C}^{2d}$ is surjective.*

16

We write

$$\Delta = \Delta_1 + \Delta_2 + \cdots + \Delta_\lambda + \cdots$$

where $\Delta_1, \Delta_2, \cdots, \Delta_\lambda, \cdots$ denote the irreducible components of Δ. Moreover we define $\mathfrak{d}_\lambda$ to be the divisor on the non-singular model of Δ_λ cut out by a general plane in W and denote by $\mathfrak{t}_\lambda$ the divisor on the non-singular model of Δ_λ composed of the points corresponding to the triple points of Δ. Now, using the theorem of Riemann-Roch, we infer readily that if the divisors $\mathfrak{d}_\lambda - \mathfrak{t}_\lambda$, $\lambda = 1, 2, 3, \cdots$, are non-special, then the above conditions α) and β) are fulfilled and therefore the cohomology group $H^1(\Delta, N)$ vanishes. Combining this with Theorem 5 we obtain the following

THEOREM 6. *If $n \geqq n_0(\Delta)$ and if the divisors $\mathfrak{d}_\lambda - \mathfrak{t}_\lambda$, $\lambda = 1, 2, 3, \cdots$, defined respectively on the non-singular models of Δ_λ, are non-special, then the surface S is regular in W.*

We denote by π_λ the genus of the non-singular model of Δ_λ, by π the virtual genus of Δ and by τ the number of triple points of Δ. Moreover we denote by τ_λ and d_λ the respective degrees of the divisors $\mathfrak{t}_\lambda$ and $\mathfrak{d}_\lambda$. Note that d_λ is the order of the curve Δ_λ. We have

$$\pi = 2\tau + 1 + \sum_\lambda (\pi_\lambda - 1), \tag{30}$$

$$3\tau = \sum_\lambda \tau_\lambda. \tag{31}$$

We derive from Theorem 6 the following

COROLLARY. *If $n \geqq n_0(\Delta)$ and if*

$$d_\lambda - \tau_\lambda > 2\pi_\lambda - 2, \tag{32}$$

then the surface S is regular in W.

Assuming that Δ is the double curve of a surface S_1 of order n_1 in W with ordinary singularities only, we shall estimate the integer $n_0(\Delta)$. We construct a non-singular model $\tilde{S}_1$ of S_1 in an obvious manner and denote the natural holomorphic map of $\tilde{S}_1$ onto S_1 by h_1 (see § 2). Moreover we denote by $\tilde{E}_1$ the complex line bundle over $\tilde{S}_1$ induced from E by the map h_1. Let f_2 be a homogeneous polynomial of degree n_2 in w which vanishes on Δ. Considering f_2 as a holomorphic section of n_2E, we denote by $h_1^*f_2$ the holomorphic section of $n_2\tilde{E}_1$ over $\tilde{S}_1$ induced from f_2 by the map h_1. Obviously the divisor $(h_1^*f_2)$ of $h_1^*f_2$ has the form

$$(h_1^*f_2) = \tilde{\Delta}_1 + \text{non-negative divisor}, \qquad \tilde{\Delta}_1 = h_1^{-1}(\Delta).$$

THEOREM 7. *If there exists a homogeneous polynomial f_2 of degree n_2 in w which vanishes on Δ such that the divisor $(h_1{}^*f_2)$ has no multiple component, then the inequality*

$$n_0(\Delta) \leqq n_1 + n_2 - 3 \tag{33}$$

holds.

Proof. Let $\Omega(kE - S_1)$ denote the sheaf over W of germs of those holomorphic sections of kE which vanish on the surface S_1. We infer readily that $\Omega(kE - S_1)$ is a subsheaf of $\Omega(kE - 2\Delta)$, while we have the isomorphism

$$\Omega(kE - S_1) \cong \Omega(k_1E), \qquad k_1 = k - n_1.$$

Hence, denoting by r_1 the restriction map to the surface S_1, we obtain the exact sequence

$$0 \to \Omega(k_1E) \to \Omega(kE - 2\Delta) \to r_1\Omega(kE - 2\Delta) \to 0. \tag{34}$$

Our purpose is to show that

$$H^\nu(W, \Omega(kE - 2\Delta)) = 0, \qquad \text{for } k \geqq n_1 + n_2 - 3, \nu = 1, 2.$$

Since the cohomology group $H^\nu(W, \Omega(k_1E))$ vanishes for $\nu = 1, 2$, we infer from (34) that, for $\nu = 1, 2$, the vanishing of $H^\nu(S_1, r_1\Omega(kE - 2\Delta))$ implies the vanishing of $H^\nu(W, \Omega(kE - 2\Delta))$. On the other hand, examining the stalks of the sheaf $r_1\Omega(kE - 2\Delta)$, we obtain the isomorphism

$$H^\nu(S_1, r_1\Omega(kE - 2\Delta)) \cong H^\nu(\bar{S}_1, \Omega(k\bar{E}_1 - 2\bar{\Delta}_1), \qquad \text{for } \nu = 0, 1, 2. \tag{35}$$

Therefore it suffices for our purpose to verify that

$$H^\nu(\bar{S}_1, \Omega(k\bar{E}_1 - 2\bar{\Delta}_1)) = 0, \qquad \text{for } k \geqq n_1 + n_2 - 3, \nu = 1, 2. \tag{36}$$

By hypothesis we have

$$(h_1{}^*f_2) = \bar{\Delta}_1 + \Gamma,$$

where Γ is a *curve* on $\bar{S}_1$. It follows that

$$n_2\bar{E}_1 = [\bar{\Delta}_1 + \Gamma],$$

while, by (8), the canonical bundle of $\bar{S}_1$ is

$$\bar{K}_1 = (n_1 - 4)\bar{E}_1 - [\bar{\Delta}_1].$$

Hence we obtain

$$\Omega(k\bar{E}_1 - 2\bar{\Delta}_1) \cong \Omega(\bar{K}_1 + k_2\bar{E}_1 + \Gamma), \qquad k_2 = k - n_1 - n_2 + 4.$$

We assume that $k \geqq n_1 + n_2 - 3$ or $k_2 \geqq 1$. Take a curve C on S_1 cut out by a general surface of order k_2. We have

$$k_2 \tilde{E}_1 = [\tilde{C}], \qquad \tilde{C} = h_1^{-1}(C),$$

and consequently

$$H^\nu(\tilde{S}_1, \Omega(k\tilde{E}_1 - 2\tilde{\Delta}_1)) \cong H^\nu(\tilde{S}_1, \Omega(\tilde{K}_1 + \tilde{C} + \Gamma)).$$

Now the vanishing of $H^2(\tilde{S}_1, \Omega(\tilde{K}_1 + \tilde{C} + \Gamma))$ is obvious, while the vanishing of $H^1(\tilde{S}_1, \Omega(\tilde{K}_1 + \tilde{C} + \Gamma))$ follows from a result of [4] (see [4], Theorem 2.3). Thus we obtain (36), q.e.d.

Assume that the surface S is regular in W. Then, by Theorems 1 and 2, S belongs to an effectively parametrized maximal analytic family $\mathfrak{F}$ of surfaces S_t, $t \in M_1$, with ordinary singularities in W whose characteristic system on each surface S_t is complete. To emphasize that $\mathfrak{F}$ is effectively parametrized, we call the dimension of the parameter manifold M_1 the *number of effective parameters of the family* $\mathfrak{F}$. Obviously the number of effective parameters of $\mathfrak{F}$ is equal to the dimension of the linear space $\Lambda(S)$. From (6), (12) and (13) we get

$$\dim \Lambda(S) = 2n^2 - 4d + p_a - \gamma + 1,$$

while we have the classical formulae

$$p_a = \binom{n-1}{3} - (n-4)d + \pi - 1, \tag{37}$$

$$\gamma = (2n-8)d + 2\tau - 4\pi + 4. \tag{38}$$

Hence we obtain the following *formula for the number* μ *of effective parameters of the family* $\mathfrak{F}$:

$$\mu = \binom{n+3}{3} - (3n-8)d - 2\tau + 5\pi - 6. \tag{39}$$

5. Examples. First we shall examine the application of the above result to some classical examples of algebraic surfaces. We employ the notation of the preceding section. For brevity we write (w, x, y, z) for the homogeneous coordinates (w_0, w_1, w_2, w_3).

1) Enriques' surface is the surface S of order 6 defined by an equation of the form

$$x^2y^2z^2 + w^2x^2y^2 + w^2y^2z^2 + w^2z^2x^2 + wxyzg = 0,$$

where g denotes a general quadratic form in w, x, y, z. The double curve Δ

of S is composed of the edges $\Delta_1, \cdots, \Delta_\lambda, \cdots, \Delta_6$ of the tetrahedron formed of the coordinate planes: $w=0$, $x=0$, $y=0$, $z=0$. Hence $d_\lambda=1$, $\pi_\lambda=0$, $\tau_\lambda=2$ and thus Δ satisfies the condition (32). In order to estimate $n_0(\Delta)$ we apply Theorem 7 to $S_1=S$. The qubic form

$$f_2=a_0xyz+a_1wyz+a_2wzx+a_3wxy$$

satisfies the hypothesis of Theorem 7. Hence $n_1=6$, $n_2=3$, and therefore $n_0(\Delta)\leqq 6$. Thus we see that the surface S is regular in W. Since $d=6$, $\tau=4$, $\pi=3$, we obtain, using (39), $\mu=25$.

2) A surface S of order 6 with a cubic rational double curve Δ. (This is one of sixteen examples of algebraic surfaces quoted by M. Noether (see Noether [8], p. 525; Baker [1], p. 288)). It is obvious that Δ satifies the condition (32). We apply Theorem 7 to $S_1=S$. There is a quadric form f_2 satisfying the hypothesis of Theorem 7. Hence $n_0(\Delta)\leqq 5$. Consequently the surface S is regular in W. Using (39) we obtain $\mu=48$.

3) A surface S of order 7 with a double elliptic quartic curve Δ_1 and a double line Δ_2 not meeting this (see Noether [8], p. 525; Baker [1], p. 228). Since $d_1=4$, $\pi_1=1$, $\tau_1=0$, $d_2=1$, $\pi_2=\tau_2=0$, the double curve Δ of S satisfies the condition (32). We apply Theorem 7 to $S_1=S$. There is a qubic form f_2 which satifies the hypothesis of Theorem 7. Hence $n_1=7$, $n_2=3$, and therefore $n_0(\Delta)\leqq 7$. Consequently the surface S is regular in W. Since $d=5$, $\pi=0$ $\tau=0$, we obtain $\mu=49$.

Next, we consider surfaces with non-singular double curves. Let $\mathfrak{F}$ be an analytic family of surfaces S_t, $t\in M_1$, of order n with ordinary singularities in W having no triple points and let Δ_t, $f(w,t)$, $\mathfrak{c}_{tt}, \cdots$ have the same meaning as in §1. Note that, by hypothesis, the double curve Δ_t of each surface S_t is non-singular. Obviously the set of the double curves Δ_t, $t\in M_1$, forms an analytic family which will be denoted by $\mathfrak{f}$. Let

$$\Phi_t=\Omega(nE-\Delta_t-\sum \mathfrak{c}_t')$$

and let N_t denote the sheaf of normal vectors on Δ_t. Then, by (19), we have the exact sequence

$$0\rightarrow\Omega(nE-2\Delta_t)\rightarrow\Phi_t\rightarrow N_t\rightarrow 0.$$

Hence we obtain the exact cohohomolgy sequence

$$(40)\qquad 0\rightarrow L_n(-2\Delta_t)\rightarrow L_n(-\Delta_t-\sum\mathfrak{c}_t')\xrightarrow{\#_t} H^0(\Delta_t,N_t)\rightarrow\cdots.$$

Since the kernel of the restriction map r_t is contained in $L_n(-2\Delta_t)$, the linear map $\#_t$ induces a linear map

$$\#_t\colon \Lambda(S_t) = r_t L(-\Delta_t - \sum \mathfrak{c}_t') \to H^0(\Delta_t, N_t),$$

where we employ the same symbol $\#_t$ to indicate the induced map. We denote by $\partial\Delta_t/\partial t$ the infinitesimal displacement of Δ_t along $\partial/\partial t$ (see [6], Definition 4). We infer from what has been said in the proof of Theorem 4 that

$$\#_t \partial S_t/\partial t = -\partial\Delta_t/\partial t. \tag{41}$$

We define $m_0(\Delta)$ to be the minimum integer m_0 such that

$$H^1(W, \Omega(kE - 2\Delta)) = 0, \qquad \text{for } k \geqq m_0.$$

It follows from (40) that, if $n \geqq m_0(\Delta_t)$, the linear map $\#_t$ maps $\Lambda(S_t)$ onto $H^0(\Delta_t, N_t)$. Hence we conclude that, *if $n \geqq m_0(\Delta_t)$ and if the characteristic system of $\mathfrak{F}$ on S_t is complet, the characteristic system of $\mathfrak{f}$ on Δ_t is also complete.*

THEOREM 8. *Let S be a surface of order n in W with ordinary singularities only whose double curve Δ is non-singular. If Δ belongs to analytic family of non-singular curves in W whose characteristic system on Δ is complete and if $n \geqq m_0(\Delta)$, then the surface S belongs to a maximal analytic family $\mathfrak{F}$ of surfaces S_t, $t \in M_t$, of order n with ordinary singularities in W whose characteristic class on each surface S_t is complete.*

Proof. By hypothesis there exists an analytic family of non-singular curves Δ_s, $s \in U$, in W such that $\Delta_0 = \Delta$, where U denotes a polycylindrical domain of the center 0 in the space of several complex variables. The completeness of the characteristic system of $\mathfrak{f}$ on Δ_0 implies the completeness of the characteristic system of $\mathfrak{f}$ on Δ_s for all s with sufficiently small norms $|s|$. Hence we may assume that $\mathfrak{f}$ is an effectively parametrized family whose characteristic system on each curve Δ_s is complete.

For all s with $|s| < \epsilon$ the cohomology group $H^1(W, \Omega(nE - 2\Delta_s))$ vanishes and the dimension of the linear space $L_n(-2\Delta_s)$ is independent of s, provided that the positive number ϵ is sufficiently small. We verify this as follows. We form the quotient sheaf

$$J_s = \Omega(nE)/\Omega(nE - 2\Delta_s).$$

Then we obtain the exact sequence

$$0 \to L_n(-2\Delta_s) \to L_n \to H^0(\Delta_s, J_s) \to H^1(W, \Omega(nE - 2\Delta_s)) \to 0.$$

This implies that

$$\dim L_n(-2\Delta_s) + \dim H^0(\Delta_s, J_s) = \dim L_n + \dim H^1(W, \Omega(nE - 2\Delta_s)).$$

On the other hand the dimensions of the linear spaces $L_n(-2\Delta_s)$ and $H^0(\Delta_s, J_s)$ are upper semi-continuous functions of s and by hypothesis $H^1(W, \Omega(nE - 2\Delta_0))$ vanishes. Therefore we infer that $H^1(W, \Omega(nE - 2\Delta_s))$ vanishes for $|s| < \epsilon$ and that

$$l = \dim L_n(-2\Delta_s) \tag{42}$$

is independent of s, $|s| < \epsilon$.

Let $f^0(w) = 0$ be the irreducible equation defining the surface S. Obviously the homogeneous polynomial $f^0(w)$ belongs to $L_n(-2\Delta_0)$. In view of (42) we can choose l linearly independent homogeneous polynomials

$$g_\nu(w, s) \in L_n(-2\Delta_s), \qquad \nu = 1, 2, \cdots, l,$$

in w whose coefficients are holomorphic functions of s, $|s| < \epsilon$, such that

$$g_l(w, 0) = f^0(w).$$

Let

$$\mu = l - 1 + \dim H^0(\Delta_0, N_0),$$

where N_0 indicates the sheaf of normal vectors on Δ_0, and let t denote a point in the space of μ complex variables $t_1, t_2, \cdots, t_\mu$. We set

$$s(t) = (t_l, t_{l+1}, \cdots, t_\mu), \qquad \text{for } t = (t_1, \cdots, t_l, \cdots, t_\mu),$$

and define

$$f(w, t) = \sum_{\nu=1}^{l-1} t_\nu g_\nu(w, s(t)) + g_l(w, s(t)), \qquad \text{for } |t| < \epsilon. \tag{43}$$

Obviously $f(w, t)$ is a homogeneous polynomial in w belonging to the linear space $L_n(-2\Delta_{s(t)})$ of which the coefficients are holomorphic functions of t, $|t| < \epsilon$, and

$$f(w, 0) = f^0(w).$$

Now we define S_t to be the surface in W determined by the equation

$$f(w, t) = 0.$$

The surface S_t has ordinary singularities only and the double curve of S_t coincides with $\Delta_{s(t)}$, provided that $|t| < \epsilon_1$, where ϵ_1 denotes a sufficiently small positive number $\leqq \epsilon$, and the surface S_0 coincides with S. Moreover the collection of the surfaces S_t, $|t| < \epsilon$, forms an analytic family $\mathfrak{F}$.

Since the cohomology group $H^1(W, \Omega(nE - 2\Delta_s))$ vanishes for $|s| < \epsilon_1$, we obtain from (40) the exact sequence

$$(44) \qquad 0 \to r_t L_n(-2\Delta_{s(t)}) \to \Lambda(S_t) \xrightarrow{\#_t} H^0(\Delta_{s(t)}) \to 0, \qquad \text{for } |t| < \epsilon_1.$$

We derive from (43) and (41) the formulae

$$(45) \qquad \begin{cases} \partial S_t/\partial t_\nu = r_t g_t(\ , s(t)), & \text{for } \nu = 1, 2, \cdots, l-1, \\ \#_t \partial S_t/\partial t_\nu = -(\partial \Delta_s/\partial s_{\nu-l+1})_{s=s(t)}, & \text{for } \nu = l, \cdots, \mu. \end{cases}$$

By hypothesis the infinitesimal displacements $\partial\Delta_s/\partial s_{\nu-l+1}$, $\nu = l, l+1, \cdots, \mu$, form a base of the linear space $H^0(\Delta_s, N_s)$. Hence, combining (45) with (44), we infer that the infinitesimal displacements $\partial S_t/\partial t_\nu, \nu = 1, 2, \cdots, \mu$, form a base of the linear space $\Lambda(S_t)$. Thus we conclude that the family $\mathfrak{F}$ is effectively parametrized and that the characteristic system of $\mathfrak{F}$ on each surface S_t is complete.

To prove that the family $\mathfrak{F}$ is maximal, we consider an analytic family $\mathfrak{F}'$ of surfaces S'_u, $u \in M'$, with ordinary singularities in W defined respectively by the equations

$$f'(w, u) = 0$$

and assume that

$$f'(w, 0) = f(w, t),$$

where M' denotes a polycylindrical domain of the center 0 in the space of several complex variables. Since, by hypothesis, the characteristic system of $\mathfrak{f}$ on each curve Δ_s is complete, the family $\mathfrak{f}$ is maximal (see [6], Theorem 2). Let Δ'_u denote the double curve of S'_u. The curve Δ'_0 coincides with $\Delta_{s(t)}$. Hence there exists a holomorphic map $h: u \to s = h(u)$ defined for u with sufficiently small $|u|$ such that $\Delta'_u = \Delta_{h(u)}$. It follows that the homogeneous polynomial $f'(w, u)$ belongs to $L_n(-2\Delta_{h(u)})$ and therefore

$$f'(w, u) = \sum_{\nu=1}^{l} a_\nu(u) g_\nu(w, h(u)),$$

where the coefficients $a_\nu(u)$ are holomorphic functions of u and $a_l(0) = 1$. Consequently, setting

$$t(u) = (a_1(u)/a_l(u), \cdots, a_{l-1}(u)/a_l(u), h_1(u), \cdots, h_{\mu-l+1}(u)),$$

we obtain

$$f'(w, u) = a_l(u) f(w, t(u)).$$

This proves that S'_u coincides with $S_{t(u)}$, provided that $|u|$ is sufficiently small, q. e. d.

4) Surfaces whose double curves are complete intersections. We fix positive integers n_1 and $n_2 \leqq n_1$ and define $\mathfrak{f}$ to be the set of all irreducible non-singular curves which are the complete intersections $S_1 \cdot S_2$ of non-singular surfaces S_1 and S_2 of respective orders n_1 and n_2 in W. We infer readily that the set $\mathfrak{f}$ forms an effectively parametrized analytic family of curves in W and that the number ν of effective parameters of $\mathfrak{f}$ is given by the formula

$$\nu = C(n_1) + C(n_2) - C(n_1 - n_2) - \delta_{n_1 n_2} - 2, \tag{46}$$

where $\delta_{n_1 n_2}$ is Kronecker's delta and

$$C(m) = \tfrac{1}{6}(m+3)(m+2)(m+1).$$

Consider a curve $\Delta = S_1 S_2 \in \mathfrak{f}$ and denote by N the sheaf of normal vectors on Δ. By a simple calculation we get

$$\dim H^0(\Delta, N) = C(n_1) + C(n_2) - C(n_1 - n_2) - \delta_{n_1 n_2} - 2, \tag{47}$$

$$\dim H^1(\Delta, N) = C(n_1 - 4) + C(n_2 - 4) - C(n_1 - n_2 - 4) - \delta_{n_1 n_2}. \tag{48}$$

Comparing (47) with (46) we infer that *the characteristic system of* $\mathfrak{f}$ *on* Δ *is complete.* To estimate the integers $m_0(\Delta)$ and $n_0(\Delta)$ we consider the exact sequence

$$0 \to \Omega(hE) \to \Omega(kE - 2\Delta) \to r_1\Omega(kE - 2\Delta) + r_2\Omega(kE - 2\Delta) \to 0, \tag{49}$$

where r_1 and r_2 denote respectively the restriction maps to the surfaces S_1 and S_2 and $h = k - n_1 - n_2$. We have

$$r_1\Omega(kE - 2\Delta) \cong r_1\Omega(k_1 E), \qquad k_1 = k - 2n_2,$$

while the exact sequence

$$0 \to \Omega([k - n_1]E) \to \Omega(kE) \to r_1\Omega(kE) \to 0$$

implies that

$$H^1(S_1, r_1\Omega(kE)) = 0, \qquad \text{for all integers } k,$$
$$H^2(S_1, r_1\Omega(kE)) = 0, \qquad \text{for } k > n_1 - 4.$$

Hence we infer from (49) that

$$H^1(W, \Omega(kE - 2\Delta)) = 0, \qquad \text{for all integers } k,$$
$$H^2(W, \Omega(kE - 2\Delta)) = 0, \qquad \text{for } k \geqq 2n_1 + n_2 - 3.$$

Thus we obtain

$$m_0(\Delta) = -\infty, \qquad n_0(\Delta) \leqq 2n_1 + n_2 - 3. \tag{50}$$

For any curve $\Delta \in \mathfrak{f}$ and for any integer $n \geqq 2n_1 + 1$ there exist surfaces of order n in W with ordinary singularities only whose double curves coincide with Δ. In fact, the curve Δ is defined by simultaneous equations

$$g(w) = h(w) = 0,$$

where $g(w)$ and $h(w)$ are homogeneous polynomials of respective degrees n_1 and n_2. Hence, for any general homogeneous polynomials $a(w)$, $b(w)$ and $c(w)$ of respective degrees $n - 2n_1$, $n - n_1 - n_2$, $n - 2n_2$, the equation

$$a(w)g^2(w) + b(w)g(w)\,h(w) + c(w)h^2(w) = 0$$

defines a surface S of order n in W which is non-singular outside Δ and has ordinary singularities on Δ. The formula (48) shows that $H^1(\Delta, N)$ vanishes if and only if $n_1 \leqq 3$, while, by (50), $n_0(\Delta) \leqq 2n_1 + n_2 - 3$. Hence, by Theorem 5, *the surface S is not regular in W if $n_1 \geqq 4$ and if $n \geqq 2n_1 + n_2 - 3$.* On the other hand, since $m_0(\Delta) = -\infty$, it follows from Theorem 8 that *the surface S belongs to a maximal family $\mathfrak{F}$ of surfaces with ordinary singularities in W whose characteristic system on S is complete.*

Apparently the above example indicates that the requirement of semi-regularity imposes a strong restriction on surfaces with ordinary singularities.

Now we shall examine the application of Theorem 1 to a regular canonical surface of genus 4 and of order 7 (see Enriques [2], pp. 273-280; Maxwell [7], pp. 306-308). Since this surface has several singular points which may be considered as degenerated forms of ordinary double points, we shall show first that *Theorem 1 can be applied to surfaces with some degenerated forms of ordinary singularities.*

By a *quasi-ordinary multiple point* of a surface S in W we shall mean a singular point p of S such that, in a neighborhood of p, S is composed of several non-singular sheets and has no non-ordinary singular point except p. By a *quasi-ordinary singular point* of S we shall mean a point which is either a quasi-ordinary multiple point or an (ordinary) cuspidal point of S. Note that ordinary double points and triple points are special cases of quasi-ordinary multiple points.

Let S be an algebraic surface of order n in W with quasi-ordinary singularities only defined by an irreducible equation $f^0(w) = 0$ and let Δ denote the *double curve,* i.e., the singular locus of S. We cover Δ by a finite number of sufficiently small neighborhoods W_i, $i \in I_\Delta$, and choose a local coordinate (x_i, y_i, z_i) on each neighborhood W_i such that $f^0_i(w)$ coincides with one of the functions

$$y_i z_i, \quad x_i y_i^2 - 4z_i^2, \quad e^0_i \prod_{\nu=1}^{m_i} (z_i - g_{i\nu}(x_i, y_i)),$$

where e^0_i denotes a non-vanishing holomorphic function of x_i, y_i, z_i and the $g_{i\nu}(x_i, y_i)$ are holomorphic functions of x_i, y_i such that $g_{i\nu}(0,0)=0$. Thus we have three divisions I_d, I_c, I_m of I_Δ and

$$(51)\qquad f^0_i(w)=f^0(w/w_{\alpha(i)})=\begin{cases} y_iz_i & \text{for } i\in I_d,\\ x_iy_i^2-4z_i^2, & \text{for } i\in I_c,\\ e^0_i\prod_{\nu=1}^{m_i}(z_i-g_{i\nu}(x_i,y_i)), & \text{for } i\in I_m.\end{cases}$$

We assume moreover that each neighborhood W_i is the polycylinder defined by the inequalities: $|x_i|<1$, $|y_i|<1$, $|z_i|<1$ and that, for $i\in I_m$, the center: $x_i=y_i=z_i=0$ of W_i is *not* an ordinary double point. For each index $i\in I_c$ we denote by $\mathfrak{c}_i$ the center of W_i which is a cuspidal point of S. Defining the linear space $L_n(-\Delta)$, the sheaf $\Phi,\cdots$ in the same manner as in § 2, we infer readily that *the surface S is regular in W in the sense that $H^1(W,\Phi)$ vanishes if and only if the linear equations $(\partial\phi_i/\partial y_i)(\mathfrak{c}_i)=0$, $i\in I_c$, are linearly independent on $L_n(-\Delta)$.* Thus Theorem 2 holds for surfaces with quasi-ordinary singularities only.

By an *analytic family of surfaces with quasi-ordinary singularities in W* we shall mean a family of surfaces S_t, $t\in M$, in W satisfying the following conditions:

i) Each surface S_t has quasi-ordinary singularities only.

ii) There is an irreducible homogeneous polynomial $f(w,t)$ in w_0, w_1, w_2, w_3 whose coefficients are holomorphic functions of $t\in M_1$ such that the equation $f(w,t)=0$ defines the surface S_t.

iii) Take an arbitrary point $t_0\in M_1$ and determine W_i, x_i, y_i, z_i, $g_{i\nu}(x_i,y_i),\cdots$ in the above manner with respect to the surface S_{t_0}. Then we find holomorphic functions $X_i(t)$, $Y_i(t)$, $Z_i(t)$, $g_{i\nu}(x_i,y_i,t)$ of x_i, y_i, z_i, t, $|x_i|<1$, $|y_i|<1$, $|z_i|<1$, $|t-t_0|<\epsilon$, satisfying the boundary conditions $X_i(t_0)=x_i$, $Y_i(t_0)=y_i$, $Z_i(t_0)=z_i$, such that

$$(52)\qquad f_i(w,t)=\begin{cases} Y_i(t)Z_i(t), & \text{for } i\in I_d,\\ X_i(t)Y_i(t)^2-4Z_i(t)^2, & \text{for } i\in I_c,\\ e_i(t)\prod_{\nu=1}^{m_i}(z_i-g_{i\nu}(x_i,y_i,t)), & \text{for } i\in I_m,\end{cases}$$

where ϵ is a positive number and $e_i(t)$ denotes a non-vanishing holomorphic function of x_i, y_i, z_i, t.

For any analytic family $\mathfrak{F}$ of surfaces S_t, $t\in M_1$, of order n with quasi-

ordinary singularities in W we define in the same manner as in § 1 the linear spaces $L_n(-\Delta_t - \sum c_i')$ and $\Delta(S_t)$, the infinitesimal displacement $\partial S_t/\partial t$, the characteristic system of $\mathfrak{F}$ on $S_t, \cdots$.

We maintain that, *if a surface S with quasi-ordinary singularities only is regular in W, then S belongs to an effectively parametrized maximal analytic family $\mathfrak{F}$ of surfaces S_t, $t \in M_1$, with quasi-ordinary singularities in W whose characteristic system on each surface S_t is complete.* Thus Theorem 1 holds for surfaces with quasi-ordinary singularities. In fact, it is a matter of triviality to verify that the construction of analytic families expounded in [5] and the derivation of Theorem 1 from Theorem 1* in § 2 can be applied to surfaces with quasi-ordinary singularities only.

Now we consider a surface S in W of order 7 defined by an equation of the form

$$w_1w_2w_3g^2 + (w_2w_3 + w_3w_1 + w_1w_2)^2h = 0,$$

where g and h denote respectively general quadric and cubic form in w_0, w_1, w_2, w_3 (see Maxwell [7], p. 308). In what follows we write (w, x, y, z) for (w_0, w_1, w_2, w_3) and set

$$f^0 = f^0(w, x, y, z) = xyzg^2 + Q^2h, \qquad Q = yz + zx + xy. \tag{53}$$

The singular locus Δ of S is composed of three lines Δ_1, Δ_2, Δ_3 and a quartic curve Δ_4 defined respectively by the equations $y = z = 0$, $z = x = 0$, $x = y = 0$ and $g = Q = 0$. The lines Δ_1, Δ_2, Δ_3 meet at the origin $O = (1, 0, 0, 0)$ and the quartic curve Δ_4 intersects each line Δ_λ transversally at two points $b_{\lambda 1}$ and $b_{\lambda 2}$. Thus the curve Δ has one triple point O and six double points b_{11}, $b_{12}, \cdots, b_{32}$. The origin O is a triple point of S.

We write

$$f^0 = Ay^2 + Byz + Cz^2, \tag{54}$$

where

$$A = C = x^2h, \quad B = xg^2 + (2x^2 + 2xy + 2xz + yz)h.$$

We infer from (54) that a point p on the line Δ_1: $y = z = 0$ is an ordinary double point of S unless the discriminant $B^2 - 4AC$ vanishes at p. The zeros of $B^2 - 4AC$ on Δ_1 are O, b_{11}, b_{12} and four points c_1, c_2, c_3, c_4 determined by the simultaneous equations

$$y = z = g^2 + 4xh = 0. \tag{55}$$

For each point b_{1k} we find a local coordinate (x_{1k}, y_{1k}, z_{1k}) of the center b_{1k} such that

$$f^0(1, x, y, z) = (z_{1k} - x_{1k}y_{1k})(z_{1k} + x_{1k}y_{1k}). \tag{56}$$

In fact, setting $w=1$ and writing

$$R=xy-xz-yz, \qquad U=4x+4y,$$

we get

$$f^0(1,x,y,z)=(h+U^{-1}g^2)Q^2-U^{-1}R^2g^2.$$

Hence, letting

(57) $$x_{1k}=g, \qquad y_{1k}=U^{-\frac{1}{2}}R, \qquad z_{1k}=(h+U^{-1}g^2)^{\frac{1}{2}}Q,$$

we obtain (56). It follows from (56) that b_{1k} is a quasi-ordinary double point of S. For each point $\mathfrak{c}_i$ we define

(58) $$x_i=B^2-4AC, \qquad y_i=z, \qquad z_i=Ay+\tfrac{1}{2}Bz,$$

where we set $w=1$. Obviously (x_i,y_i,z_i) forms a local coordinate of the center $\mathfrak{c}_i$ and

$$-4Af^0(1,x,y,z)=x_iy_i^2-4z_i^2.$$

Thus we see that S has twelve cuspidal points on three lines Δ_1, Δ_2, Δ_3, which we denote by $\mathfrak{c}_1,\mathfrak{c}_2,\cdots,\mathfrak{c}_{12}$.

We infer from (53) that a point $p\neq b_{\lambda k}$ on the quartic curve Δ_4: $g=Q=0$ is an ordinary double point on S unless h vanishes at p and that the equation $h=0$ determines twelve points $\mathfrak{c}_{13},\mathfrak{c}_{14},\cdots,\mathfrak{c}_{24}$ on Δ_4 which are cuspidal points of S. In fact, setting $w=1$ and defining

(59) $$x_i=-4xyzh, \qquad y_i=Q, \qquad x_i=xyzg,$$

we obtain

$$-4xyzf^0(1,x,y,z)=x_iy_i^2-4z_i^2.$$

Thus we conclude that S has quasi-ordinary singularities only.

The surface S is regular, i.e., the irregularity of S vanishes and the genus of S is 4. Moreover S is canonical in the sense that a general plane cuts out a canonical curve on S.

Now we prove that the linear equations

(60) $$(\partial\phi_i/\partial y_i)(\mathfrak{c}_i)=0, \qquad i=1,2,3,\cdots,24,$$

are linearly independent on the linear space $L_7(-\Delta)$, where $\phi_i=\phi(1,x,y,z)$. We infer readily that $L_7(-\Delta)$ consists of all homogeneous polynomials of the form

$$\psi Q+(\xi yz+\eta zx+\zeta xy)g, \qquad \psi\in L_5, \xi,\eta,\zeta\in L_3.$$

Setting

$$\phi=\psi Q+(\xi yz+\eta zx+\zeta xy)g$$

and using (55), (58) and (59) we obtain

$$(\partial\phi_i/\partial y_i)(c_i) = (\eta xg + \zeta xg + 2\psi x)(c_i), \quad \text{for } i = 1, 2, 3, 4,$$
$$(\partial\phi_i/\partial y_i)(c_i) = \psi(c_i), \quad \text{for } i = 13, 14, \cdots, 24.$$

Hence we conclude that the linear equations (60) are equivalent to the simultaneous linear equations

$$\tag{61} \begin{cases} \eta(c_i) + \zeta(c_i) + a_i\psi(c_i) = 0, & i = 1, 2, \cdots, 4, \\ \zeta(c_i) + \xi(c_i) + a_i\psi(c_i) = 0, & i = 5, \cdots, 8, \\ \xi(c_i) + \eta(c_i) + a_i\psi(c_i) = 0, & i = 9, \cdots, 12, \\ \psi(c_i) = 0, & i = 13, 14, \cdots, 24, \end{cases}$$

where $a_i = 2/g(c_i)$. Let r_4 denote the restriction map to the curve Δ_4. Obviously the equations

$$\tag{62} \psi(c_i) = 0, \qquad i = 13, 14, \cdots, 24,$$

are linearly independent on $H^0(\Delta_4, r_4\Omega(5E))$, while r_4 maps L_5 onto $H^0(\Delta_4, r_4\Omega(5E))$. Hence the equations (62) are linearly independent on L_5. It follows that the simultaneous equations (61) are linearly independent. This proves the linear independence of the linear equations (60).

Thus the surface S satisfies the hypothesis of Theorem 1. Consequently there exists a maximal analytic family $\mathfrak{F}$ of surfaces S_t, $t \in M_1$, with quasi-ordinary singularities in W such that S_0 coincides with S and such that, for each t, the map $\sigma_t: \partial/\partial t \to \partial S_t/\partial t$ maps the tangent space $T_t(M_1)$ isomorphically onto $\Lambda(S_t)$. Each surface S_t is regular, canonical, of genus 4 and of order 7.

We now show that the double curve Δ_t of any general member S_t of $\mathfrak{F}$ is an irreducible curve of order 7 with only one singular point (which is the triple point of S_t). It is obvious that, for any t with small norm $|t|$, the double curve Δ_t has one and only one triple point O_t in a neighborhood of O. Consider general solutions ξ, η, ζ, ψ of the simultaneous equations (61) and let

$$\phi = \psi Q + (\xi yz + \eta zx + \zeta xy) g.$$

Since ϕ belongs to $L_7(-\Delta - \sum c')$, the restriction $r\phi$ is an element of $\Lambda(S)$. Hence there exists an analytic curve $t(u)$ in M_1 with $t(0) = 0$ such that

$$\tag{63} (\partial S_{t(u)}/\partial u)_{u=0} = r\phi,$$

where u denotes a complex parameter. Now assuming that $0 < |u| < \epsilon$ where ϵ is a small positive number, we prove that the double curve $\Delta_{t(u)}$ has

no singular point except $O_{t(u)}$. For this purpose it suffices to verify that $\Delta_{t(u)}$ has no singular point in a neighborhood of b_{11}. Let

$$f(w,x,y,z,u)=0$$

be the equation of $S_{t(u)}$. In view of (63) we may assume that

$$\phi_t=\phi(1,x,y,z)=(\partial f(1,x,y,z,u)/\partial u)_{u=0}. \tag{64}$$

Since, by (56),

$$f(1,x,y,z,0)=f^0(1,x,y,z)=(z_{11}-x_{11}y_{11})(z_{11}+x_{11}y_{11}),$$

we have power series

$$Y(u)=z_{11}-x_{11}y_{11}+uY_1+u^2Y_2+\cdots,$$
$$Z(u)=z_{11}-x_{11}y_{11}+uZ_1+u^2Z_2+\cdots.$$

in u whose coefficients are holomorphic functions of x_{11}, y_{11}, z_{11} such that

$$f(1,x,y,z,u)=Y(u)Z(u). \tag{65}$$

To prove that $\Delta_{t(u)}$ has no singular point in a neighborhood of b_{11} it suffices therefore to show that Y_1-Z_1 does not vanish at b_{11}. From (64) and (65) we get

$$\phi_t=(Y_1-Z_1)x_{11}y_{11}+(Y_1+Z_1)z_{11}$$

while, by (57),

$$\phi_t=(x\zeta-x\eta+y\zeta-y\xi)2U^{-\frac{1}{2}}x_{11}y_{11}+(\cdots)z_{11}.$$

Hence we infer that

$$(Y_1-Z_1)(b_{11})=2xU^{-\frac{1}{2}}(\zeta-\eta)(b_{11})\neq 0.$$

Thus we conclude that the *general member S_t of $\mathfrak{F}$ is a regular canonical surface of genus 4 and of order 7 with ordinary singularities only whose double curve Δ_t is an irreducible curve of genus 4 of order 7 with one triple point O_t* (compare Maxwell [7], p. 307).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] H. F. Baker, *Principles of Geometry*, vol. 6. *Introduction to the Theory of Algebraic Surfaces and Higher Loci*, Cambridge University Press, 1933.

[2] F. Enriques, *Le Superficie Algebriche*, Bologna, 1949.

[3] K. Kodaira, "On a differential-geometric method in the theory of analytic stacks," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 39 (1953), pp. 1268-1273.

[4] ———, "On compact complex analytic surfaces, I," *Annals of Mathematics*, vol. 71 (1960), pp. 111-152.

[5] ———, "A theorem of completeness for analytic systems of surfaces with ordinary singularities," *Annals of Mathematics*, vol. 74 (1961), pp. 591-627.

[6] ———, "A theorem of completeness of characteristic systems for analytic families of compact submanifolds of complex manifolds," *Annals of Mathematics*, vol. 75 (1962), pp. 146-162.

[7] E. A. Maxwell, "Regular canonical surfaces of genus three and four," *Proceedings of the Cambridge Philosophical Society*, vol. 23 (1937), pp. 306-310.

[8] M. Noether, "Zur Theorie des eindeutigen Entsprechens algebraischer Gebilde, II," *Mathematische Annalen*, vol. 8, pp. 495-533.

[9] O. Zariski, *Algebraic Surfaces*, Berlin, 1935.

AN EXTENDED TORELLI THEOREM.

By HENRIK H. MARTENS.[1]

Let X be a complete non-singular curve of genus g. Let $\phi: X \to J(X)$ be a canonical map of X into its jacobian variety, normalized such that $\phi(P)=0$ for some $P \in X$. If $D=\sum d_i Q_i$ is a divisor on X, we define $\phi(D)=\sum d_i \phi(Q_i)$ and we denote by W^r the image of the set of positive divisors of degree $\leqq r$ under ϕ. We set $W^0=\{0\}$.

Torelli's theorem (see [2]) asserts that W^1 is determined up to a translation and a reflection by $J(X)$ and W^{g-1}. The object of the present paper is to establish the following extended Torelli theorem:

THEOREM. *Let X and Y be complete non-singular curves of the same genus g, and assume that $J(X)=J(Y)=J$. Denote by W^r (respectively V^r) the image of the set of positive divisors of degree $\leqq r$ on X (respectively Y) under a (normalized) canonical map. If, for some t, $1 \leqq t \leqq g-1$, W^t coincides with a translate of V^t, then there exists an automorphism λ, of J such that $\lambda(W^1)$ is a translate of V^1.*

As a corollary of this result one can get a generalization of Andreotti's version of Torelli's theorem ([1]), to the effect that the birational equivalence class of X is determined by the birational equivalence class of the symmetric product $X^{(t)}$, for any t, $1 \leqq t \leqq g-1$.

My original proof of the extended Torelli theorem was restricted to the classical case. I am indebted to Professor T. Matsusaka for pointing out to me how an abstract proof could be obtained, but any errors in the following presentation should be blamed on me.

We begin by reviewing some combinatorial results which were inspired by Hilfssätze 1 and 3 of [6].

Combinatorial preliminaries. We denote by $W^r{}_a$ the translate of W^r by an element $a \in J(X)$, and by $(W^r{}_a)^-$ the image of $W^r{}_a$ under the automorphism $u \to -u$. We set $K=\phi(Z)$, where Z is a canonical divisor on X.

Received February 4, 1963.
Revised October 16, 1964.

[1] This work has been partially supported by the National Science Foundation under grant No. NSF GP-1904.

Finally, if A and B are subsets of $J(X)$, we define

$$A \oplus B = \bigcup_{b \in B} A_b$$

$$A \ominus B = \bigcap_{b \in B} A_{-b}.$$

LEMMA 1. *Let* $0 \leqq r \leqq t \leqq g-1$. *Then*

a) $(W^{g-1}{}_a)^- = W^{g-1}{}_{-a-K}$.

b) $W^r{}_a \subset W^t{}_b \Longleftrightarrow a \in W^{t-r}{}_b$.

c) $W^t{}_b \oplus W^r{}_a = W^{t+r}{}_{a+b}$.

d) $W^t{}_b \ominus W^r{}_a = W^{t-r}{}_{b-a}$.

e) $(W^{g-1}{}_b)^- \ominus W^r{}_a = W^{g-1-r}{}_{-(a+b+K)}$.

f) $W^{g-1}{}_b \ominus (W^r{}_a)^- = (W^{g-1-r}{}_{-(a+b+K)})^-$.

g) $(W^r{}_a)^- \subset W^t{}_b \Longleftrightarrow (W^{g-1-t}{}_{-b})^- \subset W^{g-1-r}{}_{-(a+K)}$.

For proofs of (a) and (b), see [2]. (c) is a triviality, and (d) is an immediate consequence of (b) and the definition of $\ominus$ (observe that $u \in A \ominus B$ if and only if $B_u \subset A$). (e) and (f) follows from (a) and (d), and, finally, (g) is a consequence of (d) and (f), observing that if $A \subset B$, then $C \ominus B \subset C \ominus A$.

It should be observed that an inclusion of the form $(W^r{}_a)^- \subset W^t{}_b$ means that $-(a+b)$ must be the image of a positive divisor of degree $r+t$ and (projective) dimension r. Thus the existence of such an inclusion implies the existence of special series of degree $r+t$ and dimension r. With this interpretation, (g) is easily seen to be a version of the Brill-Nöther reciprocity theorem. Since the Riemann-Roch theorem is used in the proof of (b), and since (g) essentially follows from (b), (b) may be regarded as a combinatorial version of the Riemann-Roch theorem.

We remark, in passing, that with the above interpretation one easily sees that the set of linear series of degree $r+t$ and dimension r may be represented geometrically by the set $W^t \ominus (W^r)^-$. It is not hard to show that the dimension of this set must be at least $(r+1)t-rg$, provided the set is not empty. This agrees with a formula quoted by Severi ([4], p. 380) for the classical case. The proof (which we omit) is a simple dimensionality argument using Lemma 2 below.

LEMMA 2. *Let* $0 \leqq r+1 \leqq t \leqq g-1$. *Let* $x \in W^1$, $y \in W^{t-r}$. *If* $W^{r+1}{}_a \not\subset W^t{}_{a+x-y}$, *then*

$$W^{r+1}{}_a \cap W^t{}_{a+x-y} = W^r{}_{a+x} \cup A,$$

where $A = W^{r+1}{}_a \cap (W^t{}_{a-y} \ominus (W^1)^-)$.

The proof is easily adapted from that of [2] Lemma 4.

We now consider the situation assumed in the statement of our theorem.

LEMMA 3. *Let $1 \leqq t \leqq g-1$, and assume $W^t = V^t{}_c$ for some $c \in J$. Suppose for some r, $1 \leqq r \leqq t$, that $V^1 \cap W^r{}_b$ contains two distinct points u and v, (b some point in J). Then V^1 is contained in a translate of W^r or of $(W^r)^-$.*

Remark. The same conclusion would be reached if we assumed that V^1 had two distinct points in common with a translate of $(W^r)^-$.

Proof. Since u, v are in $W^r{}_b$, we have

$$W^{t-r}{}_{-b} = W^t \ominus W^r{}_b \subset V^t{}_{c-u} \cap V^t{}_{c-v}.$$

Applying Lemma 2 to the intersection on the right, we find that either $W^{t-r}{}_{-b} \subset V^{t-1}{}_c$ or else $W^{t-r}{}_{-b} \subset (V^t{}_{c-u-v} \ominus (V^1)^-)$. In the first case, $V^1 = V^t{}_c \ominus V^{t-1}{}_c \subset W^t \ominus W^{t-r}{}_{-b} = W^r{}_b$, and in the second case

$$(V^1)^- \subset V^t{}_{c-u-v} \ominus (V^t{}_{c-u-v} \ominus (V^1)^-) \subset W^t{}_{-u-v} \ominus W^{t-r}{}_{-b} = W^r{}_{b-u-v},$$

where the inclusion on the far left is an easy consequence of the definition of $\ominus$.

The extended Torelli theorem. We consider again the situation contemplated in the statement of our theorem.

If Φ is an arbitrary divisor on J, we can define an endomorphism $\alpha(W^1, \Phi)$, by setting $\alpha(W^1, \Phi)(u) = S(W^1 \cdot (\Phi_u - \Phi))$, i.e. by summing the points of the intersection with multiplicities. Let $\xi = \alpha(W^1, \Phi)$. By [5] Theorem 32, cor. 3, we have $\Phi_u - \Phi \sim W^{g-1}{}_{\xi(u)} - W^{g-1}$, whence we find

$$\alpha(V^1, \Phi) = \alpha(V^1, W^{g-1})\alpha(W^1, \Phi).$$

Observing now that $\alpha(W^1, W^{g-1}) = \alpha(V^1, V^{g-1}) = \delta$ (the identity), we get

$$\delta = \alpha(W^1, W^{g-1}) = \alpha(W^1, V^{g-1})\alpha(V^1, W^{g-1}) = \alpha(V^1, W^{g-1})\alpha(W^1, V^{g-1}).$$

Hence $\alpha(V^1, W^{g-1})$ and $\alpha(W^1, V^{g-1})$ are inverse endomorphisms, and in particular automorphisms.

We are assuming $W^t = V^t{}_c$. Let r be the smallest integer for which an inclusion of the form $V^1 \subset W^{r+1}{}_a$ or $(V^1)^- \subset W^{r+1}{}_a$ holds. Assume for simplicity that $V^1 \subset W^{r+1}{}_a$.

Clearly $r < t$. We consider the intersection $V^1 \cap W^{g-1}{}_{a+x-y}$ where $x \in W^1$, $y \in W^{g-1-r}$. It is easily shown that we can determine y such that $V^1 \not\subset W^{g-1}{}_{a+x-y}$ for almost all x (otherwise the assumption on r would be violated). This means also that $W^{r+1}{}_a \not\subset W^{g-1}{}_{a+x-y}$, and by Lemma 2 we get that for almost all $x \in W^1$,

$$V^1 \cap W^{g-1}{}_{a+x-y} = (V^1 \cap W^r{}_{a+x}) \cup (V^1 \cap A),$$

where A is independent of the choice of x. As x varies over W^1, the sum of the points in the left-hand intersection, with multiplicities, vary over a translate of the image of W^1 under the automorphism $\alpha(V^1, W^{g-1})$. But the only variable points in this intersection must come from the intersection $V^1 \cap W^r{}_{a+x}$, and this cannot contain two distinct points.

Hence the image of W^1 under $\alpha(V^1, W^{g-1})$ is contained in, and therefore identical with a translate of the set ${}^kV^1 = \{u : u = kv, v \in V^1\}$ where k is the multiplicity of the point of $V^1 \cap W^r{}_{a+x}$ in $V^1 \cdot (W^{g-1}{}_{a+x-y})$.

In the classical case it is now easily seen that V^1 and W^1 must be conformally equivalent. For the abstract case I am indebted to Professor Matsusaka for the following argument.

Let $\lambda = \alpha(V^1, W^{g-1})$. We know that $\alpha(W^1, W^{g-1}) = \delta$, and it is easily seen that $\alpha(\lambda(W^1), \lambda(W^{g-1})) = \delta$, by observing the formula

$$\lambda(W^1) \cdot ((\lambda(W^{g-1}))_u - \lambda(W^{g-1})) = \lambda(W^1 \cdot (W^{g-1}{}_{\lambda^{-1}(u)} - W^{g-1})).$$

Using Matsusaka [3] Theorem 2, and the fact that $\lambda(W^1)$ is numerically equivalent to ${}^kV^1$, we have

$$\alpha({}^kV^1, \lambda(W^{g-1})) = \alpha(\lambda(W^1), \lambda(W^{g-1})).$$

On the other hand, by Matsusaka [3] Lemma 4, we have that ${}^kV^1$ is numerically equivalent to k^2V^1. Hence

$$\alpha({}^kV^1, \lambda(W^{g-1})) = k^2\alpha(V^1, \lambda(W^{g-1})).$$

Since $\alpha(V^1, \lambda(W^{g-1}))$ is an endomorphism, $k^2 = 1$, and

$$\delta = \alpha(V^1, \lambda(W^{g-1})) = \alpha(V^1, V^{g-1}).$$

Again by Matsusaka [3] Theorem, V^{g-1} and $\lambda(W^{g-1})$ must be numerically equivalent. It is then easy to see that V^{g-1} and $\lambda(W^{g-1})$ must differ at most by a translation, and we can apply Torelli's theorem to the curves V^1 and $\lambda(W^1)$.

Had we started with the assumption $V^1 \subset (W^{r+1}{}_a)^-$, the argument would be substantially the same, and the extended Torelli theorem is established.

REFERENCES.

[1] A. Andreotti, "On a theorem of Torelli," *American Journal of Mathematics*, vol. 80 (1958), p. 801.

[2] H. Martens, "A new proof of Torelli's theorem," *Annals of Mathematics*, vol. 78 (1963), p. 107.

[3] T. Matsusaka, "On a characterization of a Jacobian variety," *Memoirs of the College of Sciences, Kyoto, Ser. A*, vol. 32 (1959).

[4] F. Severi, *Vorlesungen über Algebraische Geometrie*, B. G. Teubner, Berlin, 1921.

[5] A. Weil, *Variétés Abéliennes et courbes algébriques*, Hermann & Cie, Paris, 1948.

[6] ———, "Zum Beweis des Torellischen Satzes," *Nachrichten der Akademie der Wissenschaften in Göttingen, Mathematisch-Physikalische Klasse* (1957), Nr. 2.

ON THE EQUATIONS $a^x \pm b^x = c^y$.

By K. Szymiczek.

1. In connection with the conjecture of Catalan that the only solution x, y, z, t in natural numbers greater than one of the equation $x^z - y^t = 1$ is 3, 2, 2, 3, W. J. LeVeque [3] has proved that for given a, b the equation $a^x - b^y = 1$ has at most one solution in natural numbers x, y, except for the case $a = 3$, $b = 2$, where there are two solutions. J. W. S. Cassels [2] has given a simpler proof of that theorem. Both these papers contain also necessary conditions which allow one to determine effectively x, y.

The purpose of the present paper is to extend the results of W. J. LeVeque and J. W. S. Cassels for the case of the equations

$$a^x \pm b^x = c^y. \tag{1}$$

Namely I shall prove in Section 3 that each of these equations possesses at most one solution in natural numbers x, y (except for the trivial cases) and this solution can be effectively found.

Section 2 contains two lemmas which give necessary and sufficient conditions that $p^{\lambda+\varepsilon} \mid a^n \pm b^n$, whenever p^λ is a primitive divisor of $a^\delta + b^\delta$ and $a^\Delta - b^\Delta$, respectively.

The fact that equations (1) can have at most one solution results from the theorem of Birkhoff and Vandiver [1] on the existence of a primitive prime divisor of numbers $a^n - b^n$. Hence we obtain also a necessary condition which must be fulfilled by the number x in (1), viz. if the solution exists, then the number c has a prime factor p such that $x \mid p - 1$. In this paper I give more precise necessary conditions for solutions of (1). Note also that the proofs of the theorems presented in this paper are simpler than that of the corresponding theorem in [3].

If $p^\lambda \mid m$ and $p^{\lambda+1} \nmid m$, I shall write $p^\lambda \parallel m$. $[t_1, \cdots, t_k]$ denotes the least common multiple of the numbers $t_1, \cdots, t_k$.

2. Let a, b be natural numbers, $(a, b) = 1$, $a > b$, and let p be a prime number and $p \nmid ab$.

Received June 20, 1963.
Revised June 19, 1964.

Definition 1. An odd prime p is called a primitive divisor of $a^\Delta - b^\Delta$ (resp. $a^\delta + b^\delta$) if Δ (resp. δ) is the least natural number such that $p \mid a^\Delta - b^\Delta$ (resp. $p \mid a^\delta + b^\delta$).

$p = 2$ is called a primitive divisor of $a - b$ if $4 \mid a - b$, and of $a^2 - b^2$ if $2 \| a - b$.

In the proofs of the theorems concerning equations (1) we shall make use of the following known lemmas (cf. [1], Theorem III, and [3], Lemma 1):

LEMMA 1. *If a prime p is a primitive divisor of $a^\Delta - b^\Delta$ and $p^\lambda \| a^\Delta - b^\Delta$, then $p^{\lambda+\varepsilon} \mid a^n - b^n$ if and only if $n = \Delta p^\varepsilon q$, where q is a natural number.*

LEMMA 2. *If an odd prime p is a primitive divisor of $a^\delta + b^\delta$ and $p^\lambda \| a^\delta + b^\delta$, then $p^{\lambda+\varepsilon} \mid a^n + b^n$ if and only if $n = \delta p^\varepsilon q$, where q is an odd number.*

3. First I shall prove a theorem concerning the equation

$$a^x - b^x = c^y. \tag{2}$$

THEOREM 1. *Let a, b, c be natural numbers, $(a, b) = 1$, $a > b \geqq 1$, $c = p_1^{\alpha_1} \cdots p_k^{\alpha_k} p^\alpha$, $k \geqq 0$, where $p_1, \cdots, p_k, p$ are primes and $p_1 < \cdots < p_k < p$. Let p (resp. p_i) be a primitive divisor of $a^\Delta - b^\Delta$ (resp. $a^{\Delta_i} - b^{\Delta_i}$) and $p^\lambda \| a^\Delta - b^\Delta$ (resp. $p_i^{\lambda_i} \| a^{\Delta_i} - b^{\Delta_i}$).*

1. *If $p > 2$ and either $p_1 > 2$ or $p_1 = 2$, $\Delta_1 = 1$, then equation (2) has at most one solution: $y = \lambda/\alpha$ and the unique x is then determined by (2).*

2. *If $p > 2$ and $p_1 = 2$, $\Delta_1 = 2$ then*

(a) *either $y = \lambda/\alpha$ and x is even*

(b) *or $y = 1$ and x is odd.*

3. *If $p = 2$ and $\Delta = 1$ then equation (2) has at most one solution: $x = 1$, $y = \lambda/\alpha$.*

4. *If $p = 2$ and $\Delta = 2$ and equation (2) has a solution, then*

(a) *either has only one solution: $x = y = 1$*

(b) *or $x = 2$, $y = \lambda/\alpha$, $a = 2^{\lambda-2} + 1$, $b = 2^{\lambda-2} - 1$. If $\alpha = 1$ also $x = y = 1$ is a solution.*

Proof. Suppose that equation (2) possesses a solution in natural numbers x, y. If $p_i > 2$ then $\alpha_i y \geqq \lambda_i$, since $p_i \mid a^x - b^x$ implies $p_i^{\lambda_i} \mid a^x - b^x$.

We now consider the case $p_1 = 2$. If $\Delta_1 = 1$ then as above we get $\alpha_1 y \geqq \lambda_1$. If $\Delta_1 = 2$ and x is even, then $a^x - b^x$ is divisible by $a^2 - b^2$, and we have $\alpha_1 y \geqq \lambda_1$.

If $\Delta_1 = 2$ and x is odd, then $2 \parallel a - b$ and $2 \parallel a^x - b^x$. Hence $\alpha_1 y = 1$, and the statement 2(b) is proved. Thus excluding the case $p_1 = \Delta_1 = 2$, x odd, we may state $\alpha_i y \geqq \lambda_i$.

We shall show that $\alpha y \leqq \lambda$. So let us suppose that $\alpha y = \lambda + \epsilon$, $\epsilon > 0$ and $\alpha_i y = \lambda_i + \epsilon_i$, $\epsilon_i \geqq 0$. Since $p_i^{\lambda_i + \epsilon_i} \mid a^x - b^x$, $p^{\lambda+\epsilon} \mid a^x - b^x$, we have by Lemma 1 $\Delta_i p_i^{\epsilon_i} \mid x$ and $\Delta p^\epsilon \mid x$. It follows that

$$x = [\Delta_1 p_1^{\epsilon_1}, \cdots, \Delta_k p_k^{\epsilon_k}, \Delta p^\epsilon].$$

Putting[1] $n = [\Delta_1 p_1^{\epsilon_1}, \cdots, \Delta_k p_k^{\epsilon_k}, \Delta]$ we obtain $x = np^\epsilon$. We shall prove that

$$c^y < a^x - b^x. \tag{3}$$

We have

$$p^\epsilon < a^{n(p^\epsilon - 1)} + a^{n(p^\epsilon - 2)} b^n + \cdots + b^{n(p^\epsilon - 1)}, \tag{4}$$

since the number of terms in the above sum equals p^ϵ and $a > 1$, $\epsilon > 0$. Since $\Delta_i p_i^{\epsilon_i} \mid n$ and $\Delta \mid n$,

$$p_1^{\lambda_1 + \epsilon_1} \cdots p_k^{\lambda_k + \epsilon_k} p^\lambda \mid a^n - b^n.$$

Hence $c^y / p^\epsilon \leqq a^n - b^n$, which together with (4) gives inequality (3).

Consequently, if equation (2) has a solution, then $\alpha y \leqq \lambda$. Since, except in the case $p = \Delta = 2$, the relation $p^z \parallel a^x - b^x$ is not fulfilled by any number $z < \lambda$, we have $\alpha y = \lambda$, and hence the statements 1., 2(a), 3., result immediately.

If $p = 2$ and $\Delta = 2$, then in view of the inequality $\alpha y \leqq \lambda$ we have $x \leqq 2$, since otherwise

$$c^y = 2^{\alpha y} \leqq 2^\lambda \leqq a^2 - b^2 < a^x - b^x.$$

If $x = 1$, then in view of $2 \parallel a - b$ we get $\alpha y = 1$, hence $x = y = 1$. If $x = 2$, we consider the equation

$$a^2 - b^2 = 2^{\alpha y}. \tag{5}$$

It may be easily shown that if equation (5) has a solution a, b, y, then $a = 2^{\lambda - 2} + 1$, $b = 2^{\lambda - 2} - 1$, $\alpha y = \lambda$, where λ is a natural number $\geqq 3$. Since if $\alpha = 1$ we have also $a - b = 2$, equation (2) has two solutions and thus statement 4(b) has been proved.

From Theorem 1 we obtain the following

COROLLARY. *The number $2^y + 1$, $y > 3$, is not a power of any natural number, with an exponent greater than one.*

---

[1] It is known that $\Delta_i \mid p_i - 1$, $\Delta \mid p - 1$ and because of $p_i < p$ consequently we get $p \nmid n$.

In fact, according to Theorem 1 the equation $a^x - 1 = 2^y$ can have a solution only if $x = 1$, except for the case where $x = 2$ and $1 = 2^{y-2} - 1$, which yields $y = 3$, $a = 3$.

THEOREM 2. *Let a, b, c be natural numbers, $(a, b) = 1$, $a > b \geqq 1$. Further let c be an odd number and $c = p_1^{\alpha_1} \cdots p_k^{\alpha_k} p^{\alpha}$, $k \geqq 0$, where $p_1, \cdots, p_k, p$ are primes and $p_1 < \cdots < p_k < p$, and moreover p is a primitive divisor of $a^{\delta} + b^{\delta}$ and $p^{\lambda} \parallel a^{\delta} + b^{\delta}$. If $a > 2$, then the equation*

$$a^x + b^x = c^y \tag{6}$$

has at most one solution: $y = \lambda/\alpha$ and the corresponding x. If $a = 2$, then equation (6) has at most one solution: $y = 1$ and the corresponding x, except for the case $c = 3$, where there are exactly two solutions: $x = y = 1$ and $x = 3$, $y = 2$.

The proof of the above theorem is similar to that of Theorem 1, the second part follows from the Corollary.

The following theorem allows one to find the solutions of equation (6) in the case where c is even.

THEOREM 3. *If a, b, c are natural numbers, $(a, b) = 1$, $a > b \geqq 1$, and c is even, then equation (6) has at most one solution in natural numbers x, y.*

If $2^{\alpha} \parallel c$, $2^{\beta} \parallel a + b$, then $y = \beta/\alpha$ in the case $\alpha > 1$, and $y = 1$ or β in the case $\alpha = 1$.

Proof. Let us assume that equation (6) possesses solutions in natural numbers x, y. At first we are going to investigate the case where at least one of the numbers α, y is greater than one: $\alpha y \geqq 2$. If x were even, $x = 2x'$, then we would obtain from (6) $c^y = (a^{x'} - b^{x'})^2 + 2a^{x'}b^{x'}$, which is impossible, since according to $\alpha y \geqq 2$, we have $4 \mid c^y$, whereas the right-hand side is not divisible by 4. Consequently x is odd. Putting $c = 2^{\alpha}C$, $a + b = 2^{\beta}A$ we have

$$2^{\alpha y}C^y = (2^{\beta}A - b)^x + b^x.$$

The greatest power of the number 2 which divides the right-hand side of the above relation is 2^{β}, and hence $\alpha y = \beta$. This proves the theorem in the case where $\alpha y \geqq 2$.

Now we shall prove that if $\alpha = 1$ and equation (6) has a solution with $y = 1$, then (6) has no solution with $y > 1$. In fact, suppose that $c = a^{x_0} + b^{x_0}$ and that for some $x, y > 1$ we have

$$(a^{x_0} + b^{x_0})^y = a^x + b^x. \tag{7}$$

Similarly as in the case where $ay \geqq 2$, it can be shown that x is odd. On the other hand x_0 is even (for otherwise $a^{x_0} + b^{x_0}$ and $a^x + b^x$ would be divisible by the same power of the number 2). But we shall prove that $x_0 \mid x$.

Suppose that $x = mx_0 + r$, $0 < r < x_0$. If m is even, then $a^{mx_0} - b^{mx_0}$ is divisible by $a^{2x_0} - b^{2x_0}$, and hence by $a^{x_0} + b^{x_0}$, and since

$$a^x + b^x = a^{mx_0}(a^r + b^r) - b^r(a^{mx_0} - b^{mx_0}),$$

and (7) holds, it follows that $a^{x_0} + b^{x_0} \mid a^r + b^r$. But this is impossible, since $r < x_0$. On the other hand, if m is odd, then $a^{mx_0} + b^{mx_0}$ is divisible by $a^{x_0} + b^{x_0}$, and since

$$a^x + b^x = a^{mx_0}(a^r - b^r) + b^r(a^{mx_0} + b^{mx_0}),$$

and (7) holds, it follows that $a^{x_0} + b^{x_0} \mid a^r - b^r$, which is impossible as well.

Thus if (7) is fulfilled, then x_0 is even, x is odd and $x = mx_0$, which is contradictory. Consequently if in the case $\alpha = 1$ equation (6) has a solution with $y = 1$, then it has no solution in which $y > 1$. Thus Theorem 3 has been completely proved.

Wyższa Szkoła Pedagogiczna,
Katowice, Poland.

REFERENCES.

[1] G. D. Birkhoff and H. S. Vandiver, "On the integral divisors of $a^n - b^n$," *Annals of Mathematics*, vol. 5 (1904), pp. 173-180.

[2] J. W. S. Cassels, "On the equation $a^x - b^y = 1$," *American Journal of Mathematics*, vol. 75 (1953), pp. 159-162.

[3] W. J. LeVeque, "On the equation $a^x - b^y = 1$," *American Journal of Mathematics*, vol. 74 (1952), pp. 325-331.

ON THE IDENTITY OF WEAK AND STRONG SOLUTIONS OF DIFFERENTIAL EQUATIONS WITH LOCAL BOUNDARY CONDITIONS.[1]

By Gideon Peyser.

This paper is concerned with the identity of weak and strong solutions of linear first order (not necessarily symmetric) systems of partial differential equations without regard to type. In [1] Friedrichs showed the identity of weak and strong solutions when no boundary conditions are imposed. In [2] Friedrichs showed further that a weak solution of the Cauchy problem for a symmetric hyperbolic system is also a strong solution. Lax and Phillips [4] proved that in the regular case, and in a domain with a smooth boundary, a weak solution of a linear first order system with given boundary conditions is also a strong solution. In the singular case they proved that a weak solution is a, so called, semi-strong solution. Sarason [5] showed the equivalence of weak and strong solutions if the weak solution has a full set of weak boundary values.

In Section 1 we show that in the singular case with smooth boundary a weak solution satisfying given boundary conditions, is also a strong solution. At the same time our method provides a simple proof for the regular case. In Section 2 we consider a boundary with corners. A weak solution satisfying the given boundary conditions is then proved to be a strong solution provided the operator and the data are subject to certain additional restrictions. The exceptional corners considered by Friedrichs [3] and Lax-Phillips [4] are shown to be special cases of our permissible corners. Let

$$L \equiv \sum_{j=1}^{m} A_j \partial/\partial x_j + B, \tag{1}$$

where A_j are $n \times n$ matrix valued functions with first order continuously differentiable elements and B is an $n \times n$ matrix with continuous elements. L operates on vector valued functions $u(x) \equiv (u_1(x), \cdots, u_n(x))$, $x \equiv (x_1, \cdots, x_m)$. The domain under consideration will be denoted by D and its boundary by S.

Received September 11, 1963.

Revised received May 12, 1964.

[1] This paper was supported by the National Science Foundation research grant NSF-G21328.

At any point $x \in S$, $N(x)$ will denote a given linear subspace of the space of vector valued functions, which varies smoothly on S. The given boundary conditions are

$$u(x) \in N(x) \quad \text{for} \quad x \in S. \tag{2}$$

In the following let $\mathfrak{H}$ denote the space of vector functions, square integrable in D, with inner product

$$(u, v) = \int_D \cdots \int u \cdot v \, dx_1 \cdots dx_m$$

and norm $\| u \|^2 = (u, u)$. Throughout the paper all coefficients and functions are assumed to be real.

1. A weak solution implies a strong solution for smooth boundaries. We now assume that the boundary S belongs to the class C^2. We define the boundary matrix $\Delta(x)$, $x \in S$, by $\Delta = \sum n_j A_j$, where $(n_1, \cdots, n_m)$ are the direction cosines of the normal to the surface S. We assume in the following that Δ is possibly singular but of constant rank, say r, on and near the boundary. By this we mean the following: If we consider any small portion of the boundary and introduce a regular transformation of the independent variable x into y such that this portion of the boundary is mapped into $y_1 = 0$, then the matrix coefficient of $\partial/\partial y_1$, in the operator L, is of rank r, on and near that part of the boundary.

Our assumption is that the boundary space $N(x)$, which is spanned by smoothly varying vectors, is of constant dimension at all points $x \in S$. We require furthermore that at every point of the boundary, $N(x)$ contains the nullspace of the boundary matrix $\Delta(x)$. Let $P(x)$, $x \in S$, denote the orthogonal complement of $\Delta(x) N(x)$. We now define the adjoint boundary conditions:

$$v(x) \in P(x) \quad \text{for} \quad x \in S. \tag{3}$$

Let L^* denote the adjoint operator of L, i. e.,

$$L^* \equiv -\sum \partial/\partial x_j A_j^* + B^*,$$

where A_j^* and B^* are the transposed matrices of A_j and B.

Definition. A function u in $\mathfrak{H}$ is a weak solution of $Lu = f$, satisfying the boundary conditions (2) in the weak sense, if the relation

$$(L^* v, u) = (v, f) \tag{4}$$

holds for all smooth functions v satisfying the adjoint boundary conditions (3).

Definition. u is a strong solution of $Lu=f$ satisfying the condition (2) in the strong sense, if there exists a sequence of smooth functions $u^{(k)}$, with $\| u^{(k)}-u \| \to 0$, such that

$$\| Lu^{(k)}-f \| \to 0 \quad \text{as} \quad k\to\infty \tag{5}$$

and

$$u^{(k)}(x) \in N(x) \quad \text{for} \quad x\in S. \tag{6}$$

THEOREM 1. *A weak solution in D of $Lu=f$ satisfying the boundary conditions (2) in the weak sense, is also a strong solution satisfying the boundary conditions in the strong sense.*

We note that if the domain D is unbounded then we may assume that the weak solution u is of bounded support. This follows from the following reasoning: Let $\psi_k(x)$ be a sequence of continuously differentiable functions which together with their first order derivatives are uniformly bounded for all x and k, and such that

$$\psi_k(x)=1 \text{ for } |x|\leqq k, \qquad \psi_k(x)=0 \text{ for } |x|\geqq k+1.$$

Consider $u_k=\psi_k u$. u_k is a weak solution of an equation $Lu=f_k$, satisfying the boundary conditions (2) in the weak sense. Since $\| u-u_k \| \to 0$ and $\| f-f_k \| \to 0$ it suffices to prove Theorem 1 for the functions u_k which have bounded support.

For the proof of Theorem 1 we first localize and normalize the problem. In this we follow Lax and Phillips [4]. Let $x^{(0)}$ be any point of the boundary S of the domain D. The following transformations of the independent variables, the dependent variables and the given equation can be achieved locally for a sufficiently small neighborhood $D^{(0)}$ $(\subset D)$ of $x^{(0)}$. A smooth regular coordinate transformation of x into y will map $D^{(0)}$ into $y_1>0$ with the relevant part of the boundary, say $S^{(0)}$, mapping into $y_1=0$. Since the boundary matrix is of constant rank r, on and near S, we may now assume that in $D^{(0)}$ the matrix A_1, which is the coefficient of $\partial/\partial y_1$, is of the constant rank r. A smooth regular transformation of the dependent variable will change A_1 such that in $D^{(0)}$ it will be of the form

$$A_1=\begin{Vmatrix} A_{11} & 0 \\ A_{21} & 0 \end{Vmatrix}, \tag{7}$$

where A_{11} is an $r\times r$ non singular matrix. Furthermore from the assumption that the boundary space N contains the null-space of the boundary

matrix it follows that we can transform the first r components of the dependent variables in $D^{(0)}$ such that on $S^{(0)}$ the boundary conditions are of the form

$$u_1 = \cdots = u_p = 0, \tag{8}$$

where p ($\leqq r$) is the codimension of N. Finally, multiplication of the operator L by a suitable smooth non-singular matrix will change the operator, in $D^{(0)}$, such that A_1 is the diagonal matrix with 1's in the first r positions of the main diagonal and 0's everywhere else.

$$A_1 = \operatorname{diag}(1, \cdots, 1, 0, \cdots, 0). \tag{9}$$

To localize the given problem we cover the closure of D (or the support of u if D is unbounded) by a finite number of open spheres T_i with radius δ_i. $2T_i$ will denote the sphere concentric to T_i with radius $2\delta_i$. Γ_i and $2\Gamma_i$ are the intersections of D with T_i and $2T_i$ respectively. We may assume the radii $2\delta_i$ to be small enough such that in any set $2\Gamma_i$ that intersects the boundary the above described transformations of the variables and the given equation can be carried out. Now let $\phi_i(x)$ be a smooth partition of unity, with the support of ϕ_i in T_i, such that $\sum \phi_i(x) = 1$. We put $w_i = \phi_i u$. Then the vector w_i is a weak solution of

$$Lw_i = g_i, \tag{10}$$

where the vector $g_i = \phi_i f + (\sum_j A_j \partial\phi_i/\partial x_j) u$.

We temporarily drop the subscript i, denoting $w = w_i$ and $g = g_i$. w satisfies the boundary conditions (2) in the weak sense. The support of w and g is in Γ_i. If $2\Gamma_i$ does not intersect the boundary S then it follows from Friedrichs' Lemma [1] that w can be approximated by a sequence of smooth functions $w^{(k)}$ with support in $2\Gamma_i$ such that $\| w^{(k)} - w \| \to 0$ and $\| Lw^{(k)} - g \| \to 0$.

We now consider the case that $2\Gamma_i$ intersects S in a set which we denote by $2S_i$. We may assume that $2\Gamma_i$ lies in $y_1 > 0$ and $2S_i$ lies on $y_1 = 0$. The boundary conditions may now be assumed to be of the form (8) and the matrix A_1 of the form (9). Since w vanishes outside Γ_i we may further assume that D is the whole half-space $y_1 > 0$ and that A_1 is of the form (9) throughout this half-space. The adjoint boundary conditions are of the form

$$v_{p+1} = \cdots = v_r = 0. \tag{11}$$

We shall utilize the following notation for the independent variables

$y \equiv (y_1, \cdots, y_m) \equiv (y_1, y')$, $y' \equiv (y_2, \cdots, y_m)$. We introduce some variants of the Friedrichs mollifiers. Let $j(s)$ be an infinitely differentiable non-negative function of the single variable s with support in the interval $-1 \leq s \leq 1$ and such that $\int j(s)\,ds = 1$. $k_\epsilon(y_1 - \bar{y}_1)$ is an $n \times n$ diagonal matrix whose first p elements on the main diagonal are

$$\epsilon^{-1} j(\epsilon^{-1}(y_1 - \bar{y}_1) - 2),$$

the next $r - p$ elements are

$$\epsilon^{-1} j(\epsilon^{-1}(y_1 - \bar{y}_1) + 2),$$

and the last $n - r$ elements are

$$\epsilon^{-1} j(\epsilon^{-1}(y_1 - \bar{y}_1)).$$

K_ϵ is the integral mollifier, mollifying in the y_1 direction, operating on functions $a \in \mathfrak{H}$,

$$K_\epsilon a = \int_0^\infty k_\epsilon(y_1 - \bar{y}_1)\, a(\bar{y}_1, y')\, d\bar{y}_1, \tag{12}$$

and the adjoint operator K_ϵ^* is defined by

$$K_\epsilon^* a = \int_0^\infty k_\epsilon(\bar{y}_1 - y_1)\, a(\bar{y}_1, y')\, d\bar{y}_1. \tag{13}$$

Next, the kernel $q_\eta(y' - \bar{y}')$ is the $n \times n$ diagonal matrix all of whose elements on the main diagonal are

$$\eta^{-m+1} \prod_{\nu=2}^{m} j(\eta^{-1}(y_\nu - \bar{y}_\nu)).$$

Q_η and Q_η^* are integral mollifiers, mollifying in the y' directions:

$$Q_\eta a = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} q_\eta(y' - \bar{y}')\, a(y_1, \bar{y}')\, d\bar{y}_2 \cdots d\bar{y}_m \tag{14}$$

$$Q_\eta^* a = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} q_\eta(\bar{y}' - y')\, a(y_1, \bar{y}')\, d\bar{y}_2 \cdots d\bar{y}_m. \tag{15}$$

Finally we have the mollifiers, $R_{\epsilon\eta}$ and $R_{\epsilon\eta}^*$, mollifying in all the variables y:

$$R_{\epsilon\eta}^* a = K_\epsilon Q_\eta a, \tag{16}$$

$$R_{\epsilon\eta} a = K_\epsilon^* Q_\eta^* a. \tag{17}$$

We list the following, well known properties of these mollifiers: For a and b in $\mathfrak{H}$

$$(R_{\epsilon\eta}a, b) = (a, R_{\epsilon\eta}^*b), \tag{18}$$

$$\| K_\epsilon a - a \| \to 0, \quad \| K_\epsilon^* b - b \| \to 0, \text{ as } \epsilon \to 0, \tag{19a}$$

$$\| Q_\eta a - a \| \to 0, \quad \| Q_\eta^* b - b \| \to 0 \text{ as } \eta \to 0, \tag{19b}$$

$$\| R_{\epsilon\eta}^* a - a \| \to 0, \quad \| R_{\epsilon\eta}^* b - b \| \to 0 \text{ as } \epsilon, \eta \to 0. \tag{19c}$$

From the different displacements of the support of the elements of the kernel $k_\epsilon(y_1 - \bar{y}_1)$ it follows that the first p components of $R_{\epsilon\eta}a$ and the components from $p+1$ to r of $R_{\epsilon\eta}^*b$ vanish on $y_1 = 0$.

We now return to the weak solutions w of (10) satisfying the data (8) in the weak sense. For arbitrary $v \in \mathfrak{H}$ the function $R_{\epsilon\eta}^*v$ is smooth and satisfies the adjoint boundary conditions (11). Hence

$$(L^*R_{\epsilon\eta}^*v, w) = (R_{\epsilon\eta}^*v, g). \tag{20}$$

Therefore

$$(v, (L^*R_{\epsilon\eta}^*)^*w) = (v, R_{\epsilon\eta}g), \tag{21}$$

where $(L^*R_{\epsilon\eta}^*)^*$ is the adjoint integral operator of $L^*R_{\epsilon\eta}^*$. Since (21) holds for all v in $\mathfrak{H}$ it follows that

$$(L^*R_{\epsilon\eta}^*)^*w = R_{\epsilon\eta}g. \tag{22}$$

Our aim now is to show that there exist sequences $\epsilon_k \to 0$ and $\eta_k \to 0$ such that if ϵ and η approach zero through these sequences then

$$\| (L^*R_{\epsilon\eta}^*)^*w - LR_{\epsilon\eta}w \| \to 0. \tag{23}$$

Assume for the moment that we have proved (23). From (22) and (19c) it then follows that

$$\| LR_{\epsilon\eta}w - g \| \to 0, \qquad \| R_{\epsilon\eta}w - w \| \to 0.$$

$R_{\epsilon\eta}w$ satisfies the boundary conditions (8) in the strict sense. Therefore w is a strong solution of (10) satisfying the boundary conditions (8) in the strong sense.

We proceed to prove (23). Denote $D_j = \partial/\partial y_j$ and $\bar{D}_j = \partial/\partial \bar{y}_j$. $(L^*R_{\epsilon\eta}^*)^*$ is an integral operator with the kernel

$$-\bar{D}_1 k_\epsilon q_\eta A_1 - \sum_{j=2}^{m} \bar{D}_j k_\epsilon q_\eta A_j(\bar{y}) + k_\epsilon q_\eta B(\bar{y}).$$

We write therefore symbolically

$$(L^*R_{\epsilon\eta}^*)^* \equiv -\bar{D}_1 R_{\epsilon\eta} \bar{A}_1 \cdot - \sum_{j=2}^{m} \bar{D}_j R_{\epsilon\eta} \bar{A}_j \cdot + R_{\epsilon\eta}\bar{B},$$

where the dot denotes that the differentiation does not apply to functions on the right of the dot. We consider next the kernel:

$$[-\bar{D}_1 A_1 - \sum_{j=2}^{m} \bar{D}_j A_j(\bar{y}) + B(\bar{y})] k_\epsilon q_\eta.$$

The corresponding integral operator is symbolically denoted by

$$\bar{L}R_{\epsilon\eta} \equiv [-\bar{D}_1 A_1 - \sum_{j=2}^{m} \bar{D}_j \bar{A}_j + \bar{B}] R_{\epsilon\eta}.$$

LEMMA 1.1. *For fixed* $\eta > 0$,

$$\| (L^* R_{\epsilon\eta}{}^*)^* w - \bar{L} R_{\epsilon\eta} w \| \to 0 \text{ as } \epsilon \to 0.$$

Proof. Since A_1 is a diagonal matrix, the kernel $k_\epsilon q_\eta$ commutes with A_1. Hence $\bar{D}_1 A_1 R_{\epsilon\eta} \cdot w = \bar{D}_1 R_{\epsilon\eta} A_1 \cdot w$. Consider next $\bar{D}_j \bar{A}_j R_{\epsilon\eta} \bar{A}_j \cdot w$, $j = 2, \cdots, m$. The kernel q_η commutes with all $n \times n$ matrices. In particular it commutes with $\bar{A}'_j \equiv A_j(y_1, \bar{y}')$. From (19a) we deduce that

$$\| (\bar{D}_j \bar{A}_j R_{\epsilon\eta} - \bar{D}_j R_{\epsilon\eta} \bar{A}_j) \cdot w \| \leqq \| (\bar{D}_j \bar{A}_j R_{\epsilon\eta} - \bar{D}_j \bar{A}'_j Q_\eta) \cdot w \| + \| (\bar{D}_j R_{\epsilon\eta} \bar{A}_j - \bar{D}_j Q_\eta \bar{A}'_j) \cdot w \| \to 0 \tag{24}$$

as $\epsilon \to 0$. Similarly, q_η commutes with $\bar{B}' \equiv B(y_1, \bar{y}')$. Therefore

$$\| (\bar{B} R_{\epsilon\eta} - R_{\epsilon\eta} \bar{B}) w \| \leqq \| (\bar{B} R_{\epsilon\eta} - \bar{B}' Q_\eta) w \| + \| (R_{\epsilon\eta} \bar{B} - Q_\eta \bar{B}') w \| \to 0$$

as $\epsilon \to 0$. This proves Lemma 1.1.

We now choose a sequence $\eta_k \to 0$. From Lemma 1.1 it follows that for every η_k we can choose a corresponding ϵ_k, with $\epsilon_k \to 0$, such that if η and ϵ approach zero through η_k and ϵ_k then

$$\| (L^* R_{\epsilon\eta}{}^*)^* w - \bar{L} R_{\epsilon\eta} w \| \to 0. \tag{25}$$

We may stipulate for later purposes that $\epsilon_k \leqq \eta_k$.

LEMMA 1.2. *If* $\epsilon, \eta \to 0$ *with* $\epsilon \leqq \eta$, *then*

$$\| \bar{L} R_{\epsilon\eta} w - L R_{\epsilon\eta} w \| \to 0. \tag{26}$$

Proof. Since A_1 is a constant matrix it follows that

$$-\bar{D}_1 A_1 R_{\epsilon\eta} \cdot w = A_1 D_1 R_{\epsilon\eta} \cdot w.$$

Hence

$$\begin{aligned} \bar{L} R_{\epsilon\eta} w - L R_{\epsilon\eta} w &= -\sum_{j=2}^{m} (\bar{D}_j \bar{A}_j + A_j D_j) R_{\epsilon\eta} \cdot w + (\bar{B} - B) R_{\epsilon\eta} w \\ &= [\sum_{j=2}^{m} \bar{D}_j (A_j - \bar{A}_j) + (\bar{B} - B)] R_{\epsilon\eta} \cdot w. \end{aligned} \tag{27}$$

Since $\epsilon \leqq \eta$ and since the derivative with respect to $\bar{y}_1$ does not appear on the right hand side of (27), it follows that we can apply the basic result of Friedrichs [2, Lemma 14], from which we conclude that

$$\| [\sum_{j=2}^{m} \bar{D}_j(A_j - \bar{A}_j) + (\bar{B} - B)] R_{\epsilon\eta} \cdot w \| \to 0 \text{ as } \epsilon, \eta \to 0.$$

This completes the proof of Lemma 1.2.

From Lemmas 1.1 and 1.2 it now follows that we have sequences $\epsilon_k, \eta_k \to 0$ such that (23) holds. Hence $w = w_i$ is a strong solution of (10) satisfying the conditions (8) in the strong sense.

If we now combine the results for the sets $2\Gamma_i$ interior to D with the results for the sets intersecting the boundary, it follows that $u = \sum w_i$ is a strong solution of (1) satisfying the boundary conditions (2) in the strong sense. This completes the proof of Theorem 1.

2. Boundaries with corners. We shall first treat the case of a corner consisting of two sides. We assume that this corner can be mapped locally into two coordinate planes, say $x_1 = 0$ and $x_2 = 0$. The boundary space N, containing the nullspace of the boundary matrix Δ, is now assumed to vary smoothly and to be of constant dimensions on each side of the corner.

Definition. The boundary conditions (2) are said to be normalized at x if the boundary space $N(x)$ is of the form

$$(28) \qquad u_{\nu_1} = \cdots = u_{\nu_\gamma} = 0 \qquad \gamma \leqq n.$$

In Section 1 we showed that under the assumptions of Theorem 1, it is possible to transform locally the independent and the dependent variables, by smooth regular transformations, such that the boundary conditions are normalized and the boundary matrix is diagonalized with constant elements. We shall extend the results of Section 1 to domains with corners provided such transformations can be achieved simultaneously on both sides of the corner.

THEOREM 2. *If u is a weak solution in D of $Lu = f$, $f \in \mathfrak{H}$, satisfying the boundary conditions* (2) *in the weak sense, and if the equation and its weak solution can be localized* (*as in Section* 1) *such that for any sufficiently small local set $2\Gamma_i$ which*

i) *does not intersect a corner, the assumptions of Theorem* 1 *hold,*

ii) *intersects a corner, the set* $2\Gamma_i$ *can be mapped, by a* C^2 *mapping, into* $y_1 > 0$, $y_2 > 0$ *with the corresponding boundary patches mapping into* $y_1 = 0$ *and* $y_2 = 0$ *respectively; the dependent variable and the equation can be changed by smooth regular transformations such that the boundary conditions are normalized simultaneously on both sides of the corner, and* A_1 *and* A_2 *both transform into diagonal matrices with constant elements,*
then u *is also a strong solution satisfying* (2) *in the strong sense.*

Proof. If $2\Gamma_i$ does not intersect a corner then the proof of Theorem 1 provides the desired sequence of approximating functions satisfying the boundary conditions and vanishing outside $2\Gamma_i$. We have, therefore, to consider only the case that the set $2\Gamma_i$ intersects a corner. We may assume this set to be in the quarter space $y_1 > 0$, $y_2 > 0$ and the two boundary patches, say $2S_i^{(1)}$ and $2S_i^{(2)}$ to be on $y_1 = 0$ and $y_2 = 0$ respectively.

On $2S_i^{(1)}$ the boundary conditions have the form

$$u_{\alpha_j} = 0 \qquad j = 1, \cdots, j_0. \tag{29}$$

Since A_1 is diagonal it follows that the adjoint boundary conditions have the form

$$v_{\beta_k} = 0 \qquad k = 1, \cdots, k_0, j_0 + k_0 \leqq n, \tag{30}$$

where $\alpha_j \neq \beta_k$ for all j and k. On $2S_i^{(2)}$ the boundary conditions are

$$u_{\gamma_j} = 0 \qquad j = 1, \cdots, j_1 \tag{31}$$

and the adjoint boundary conditions, since A_2 is diagonal, are

$$v_{\delta_k} = 0 \qquad k = 1, \cdots, k_1, j_1 + k_1 \leqq n, \tag{32}$$

where $\gamma_j \neq \delta_k$ for all j and k.

We introduce a slightly different version of the mollifiers.

$$R'_{\epsilon\eta} w = \int_D \cdots \int r'_{\epsilon\eta}(y - \bar{y}) w(\bar{y}) \, d\bar{y}_1 \cdots d\bar{y}_m,$$

where $r'_{\epsilon\eta}(y - \bar{y})$ is a diagonal $n \times n$ matrix whose elements $(r'_{\epsilon\eta})_\rho$, $\rho = 1, \cdots, n$, on the main diagonal are

$$\epsilon^{-2}\eta^{-m+2} j(\epsilon^{-1}(y_1 - \bar{y}_1 - 2\epsilon_\rho^{(1)})) j(\epsilon^{-1}(y_2 - \bar{y}_2 - 2\epsilon_\rho^{(2)})) \prod_{\nu=3}^{m} j(\eta^{-1}(y_\nu - \bar{y}_\nu)),$$

where $\epsilon_\rho^{(1)}$ is equal to ϵ in the components corresponding to those of (29) and equal to $-\epsilon$ in the components corresponding to (30). Similarly $\epsilon_\rho^{(2)}$ is equal to ϵ in the components corresponding to (31) and equal to $-\epsilon$ in

the components corresponding to (32). The rest of $\epsilon_p^{(1)}$ and $\epsilon_p^{(2)}$ are zero. Denoting, as before, $w = \phi_i u$ it now follows that $R'_{\epsilon\eta}w$ are smooth functions satisfying the boundary conditions (29) and (31) and vanishing outside $2\Gamma_i$. Furthermore, for any $v \in \mathfrak{H}$, $R'_{\epsilon\eta}{}^*v$ satisfies the adjoint boundary conditions (30) and (32). Similar to the proof of Theorem 1 we now can show that there exist sequences $\epsilon, \eta \to 0$ such that

$$\| R'_{\epsilon\eta}w - w \| \to 0 \text{ and } \| LR'_{\epsilon\eta}w - g \| \to 0,$$

from which then follows the proof of Theorem 2.

The above result can be applied to the case of corners consisting of up to m sides provided the assumptions of Theorem 2 hold simultaneously on all sides of the corner.

We conclude by extending our results to two other types of corners. These include the corners treated by Friedrichs [3, § 5] and Lax-Phillips [4, § 4]. We consider two sided corners that map locally into $y_1 > 0$, $y_2 > 0$ with the two sides of the corner mapping into $y_1 = 0$ and $y_2 = 0$ respectively. On one side of the corner, say $y_2 = 0$, we assume either (i) the boundary space N is the whole space of vector functions (i. e., there are no restrictions on the boundary conditions) or (ii) the adjoint boundary space P is the whole space of vector functions (i. e., there are no restrictions on the adjoint boundary conditions).

In the case of no restrictions on the boundary conditions on $y_2 = 0$, we use the mollifier

$$R''_{\epsilon\eta}w = \int_D \cdots \int r''_{\epsilon\eta}(y - \bar{y})\, w(\bar{y})\, d\bar{y}_1 \cdots d\bar{y}_m.$$

$r''_{\epsilon\eta}(y - \bar{y})$ is the diagonal matrix equal to $k_\epsilon(y_1 - \bar{y}_1) q''_\eta(y' - \bar{y}')$, where $k_\epsilon(y_1 - \bar{y}_1)$ is the diagonal matrix in (12) and $q''_\eta(y' - \bar{y}')$ is the diagonal matrix all of whose elements on the main diagonal are

$$\eta^{-m+1} j(\eta^{-1}(y_2 - \bar{y}_2) + 2) \prod_{\nu=3}^{m} j(\eta^{-1}(y_\nu - \bar{y}_\nu)).$$

It is readily seen that $R''_{\epsilon\eta}w$ and $R''_{\epsilon\eta}{}^*v$ satisfy the boundary conditions and adjoint boundary conditions respectively. The kernel $q''_\eta(y' - \bar{y}')$ commutes with the matrices A_j, $j = 2, \cdots, m$. We can therefore apply the proof of Theorem 1, to the present case, to show that a weak solution is also a strong one. In the second case when there are no restrictions on the adjoint boundary conditions on $y_2 = 0$, we use the mollifier

$$R'''_{\epsilon\eta}w = \int_D \cdots \int r'''_{\epsilon\eta}(y - \bar{y})\, w(\bar{y})\, d\bar{y}_1 \cdots d\bar{y}_m,$$

where $r'''_{\epsilon\eta}(y-\bar{y}) = k_\epsilon(y_1-\bar{y}_1)q'''_\eta(\bar{y}'-\bar{y}')$, and $q'''_\eta(y'-\bar{y}')$ is the diagonal matrix all of whose elements on the main diagonal are

$$\eta^{-m+1}j(\eta^{-1}(y_2-\bar{y}_2)-2)\prod_{\nu=3}^{m}j(\eta^{-1}(y_\nu-\bar{y}_\nu)).$$

We can then proceed as before.

PRATT INSTITUTE.

REFERENCES.

[1] K. O. Friedrichs, "The identity of weak and strong extensions of differential operators," *Transactions of the American Mathematical Society*, vol. 55 (1944), pp. 132-151.

[2] ———, "Symetric hyperbolic linear differential equations," *Communications on Pure and Applied Mathematics*, vol. 7 (1954), pp. 345-392.

[3] ———, "Symmetric positive linear differential equations," *Communications on Pure and Applied Mathematics*, vol. 11 (1958), pp. 333-418.

[4] P. D. Lax and R. S. Phillips, "Local boundary conditions for dissipative symmetric linear differential operators," *Communications on Pure and Applied Mathematics*, vol. 13 (1960), pp. 427-455.

[5] L. Sarason, "On weak and strong solutions of boundary value problems," *Communications on Pure and Applied Mathematics*, vol. 15 (1962), pp. 237-288.

ON QUASI-UNMIXED SEMI-LOCAL RINGS AND THE ALTITUDE FORMULA.*

By Louis J. Ratliff, Jr.

1. Introduction. The terminology used in this paper is the same as that in [2]. In particular, a semi-local (Noetherian) ring R is *quasi-unmixed* in case for every minimal prime divisor q of the zero ideal in the completion of R it is true that depth $q =$ altitude R. An integral domain R satisfies the *altitude formula* in case it is true that for every finitely generated integral domain T over R, if p is a prime ideal in T, then

$$\text{altitude } T_p + \operatorname{trd}(T/p)/(R/p \cap R) = \text{altitude } R_{(p\cap R)} + \operatorname{trd}(T/R)$$

(where $\operatorname{trd}(A/B)$ denotes the transcendence degree of the integral domain A over its subdomain B).

In this note it is proved that if R is a quasi-unmixed semi-local integral domain, if T is an integral domain which is finitely generated over R, and if $p_1, \cdots, p_r$ are prime ideals in T whose heights are the same then $T_{(T-\bigcup_1^r p_i)}$ is a quasi-unmixed semi-local domain (Corollary 2.4), and satisfies the altitude formula and the second chain condition for prime ideals (Proposition 2.2 and Corollary 2.7.) (In particular, an analytically irreducible local domain satisfies the altitude formula.) It is also proved that every quasi-unmixed semi-local ring satisfies the second chain condition for prime ideals (Corollary 2.7). These results follow quite readily from Proposition 2.1, and a large part of this paper is devoted to proving a number of lemmas which lead to a proof of this proposition. A number of these lemmas are proved in [2], and when this is the case the reader is referred to [2] for the proof.

In general the methods used in this paper are quite similar to those used by Nagata in [1] to prove some results on unmixed semi-local rings.

2. Quasi-unmixed semi-local rings and the altitude formula.

Lemma 2.1. *Let R be a quasi-unmixed semi-local ring, and let p be a prime ideal in R. Then R/p is quasi-unmixed, and* height $p +$ depth p $=$ altitude R.

* This work was supported by the National Science Foundation, Grant G21659.
Received October 10, 1963.

Proof. See ([2], p. 125).

COROLLARY 2.1. *If R is a quasi-unmixed semi-local ring, then R satisfies the first chain condition for prime ideals.*

Proof. This follows immediately from Lemma 2.1 by induction on altitude R ([2], p. 125).

LEMMA 2.2. *Let R be a semi-local ring, and let $q_1, \cdots, q_t$ be the minimal prime divisors of 0. R is quasi-unmixed if and only if* depth q_i $=$ altitude R *and R/q_i is quasi-unmixed* $(i = 1, \cdots, t)$.

Proof. If R is quasi-unmixed, then R/q_i is quasi-unmixed and depth q_i $=$ altitude R $(i = 1, \cdots, t)$ by Lemma 2.1. Conversely if q^* is a minimal prime divisor of 0 in the completion R^* of R, then q^* is a minimal prime divisor of q_iR^* for some i ([2], p. 62). Hence depth $q^* =$ altitude R^*/q^* $=$ altitude $R^*/q_iR^* =$ altitude $R/q_i =$ depth $q_i =$ altitude $R =$ altitude R^*. Therefore R is quasi-unmixed, q. e. d.

COROLLARY 2.2. *Let R be a quasi-unmixed semi-local ring and let A be an ideal in R. R/A is quasi-unmixed if and only if* height $A =$ altitude A.

Proof. Let $p_1, \cdots, p_t$ be the minimal prime divisors of A. Then R/p_i is quasi-unmixed (Lemma 2.1), so $R/A/p_i/A$ is quasi-unmixed, hence R/A is quasi-unmixed if and only depth $p_i/A =$ altitude R/A $(i = 1, \cdots, t)$ (Lemma 2.2), and this is so if and only if the p_i have the same depth. Since depth $p_i +$ height $p_i =$ altitude R (Lemma 2.1), the depths of the p_i are equal if and only if their heights are equal, hence if and only if height $A =$ altitude A, q. e. d.

LEMMA 2.3. *Let R be a complete quasi-unmixed semi-local ring, let $p_1, \cdots, p_r$ be prime ideals in R whose heights are the same, and let S be the complement in R of the union of the p_i. Then R_S is quasi-unmixed.*

Proof. Since R is a complete quasi-unmixed semi-local ring, R satisfies the first chain condition for prime ideals (Corollary 2.1). Hence if q is a minimal prime divisor of 0 which is contained in some p_i, then depth qR_S $=$ altitude $R_S/qR_S =$ altitude $(R/q)_{p_i/q} =$ height $p_i/q =$ height $p_i =$ altitude R_S. Also $R_S/qR_S = (R/q)_{(S+q)/q}$ and R/q is a complete local domain. Hence R_S/qR_S is a homomorphic image of a regular local ring, since R/q is. Therefore R_S/qR_S is unmixed ([2], p. 126). Hence by Lemma 2.2 R_S is quasi-unmixed, q. e. d.

LEMMA 2.4. *Let* $(R; M_1, \cdots, M_h)$ *and* $(R^*; M^*_1, \cdots, M^*_h)$ *be semi-local rings such that:* (1) height $M_i =$ altitude $R =$ altitude $R^* =$ height M^*_i $(i = 1, \cdots, h)$; (2) $M_i R^*$ *is* M^*_i*-primary* $(i = 1, \cdots, h)$; (3) R *is a subspace of* R^*; *and* (4) R^* *is quasi-unmixed. Then* R *is quasi-unmixed.*

Proof. Let T and T^* be the completions of R and R^* respectively, and let q be a minimal prime divisor of 0 in T. Let altitude $R = n$. If $n = 0$, then R is quasi-unmixed. Hence, it may be assumed that $n > 0$. By (3) T is a subring of T^*, hence the radical of T^* contracts in T to the radical of T. By the uniqueness of the primary decomposition of the radical of T there exists a minmal prime divisor q^* of the 0 in T^* such that $q^* \cap T = q$. Let N^* be the maximal ideal in T^* which contains q^*, and let $N = N^* \cap T$. Then N is the maximal ideal in T which contains q, and $n =$ height N^* $=$ height N by (1). Say $N^* \cap R = M_1$. Let p^*_i $(i = 1, \cdots, n)$ be a minimal prime divisor of $(p^*_{i-1}, a_i)T^*$, where $p^*_0 = q^*$ and a_i is an element in M_1 which is not in $p^*_{i-1} \cap R$. By (2) the elements a_i exist since by construction and by (4) height $p^*_i = i$ (Lemma 2.1 and the Principal Ideal Theorem). By construction $p^*_i \cap T \supset p^*_{i-1} \cap T$, hence depth $q = n =$ altitude T. Therefore R is quasi-unmixed, q. e. d.

LEMMA 2.5. *Let* R *be a quasi-unmixed semi-local ring, let* $p_1, \cdots, p_r$ *be prime ideals in* R *whose heights are the same, and let* S *be the complement in* R *of the union of the* p_i. *Then* R_S *is quasi-unmixed.*

Proof. Let R^* be the completion of R, and let p^*_i be a minimal prime divisor of $p_i R^*$ $(i = 1, \cdots, r)$. Since $\otimes_R R^*$ is exact ([2], p. 57), the theorem of transition holds for R and R^* ([2], p. 64), hence height p^*_i $=$ height p_i ([2], p. 75). Let S^* be the complement in R^* of the union of the p^*_i. Since $p^*_i \cap R = p_i$ ([3], p. 269), the theorem of transition holds for R_S and $R^*_{S^*}$ and R_S is a subspace of $R^*_{S^*}$ ([2], p. 65). Hence by Lemmas 2.3 and 2.4 R_S is quasi-unmixed, q. e. d.

LEMMA 2.6. *Let* $(R; M_1, \cdots, M_h)$ *be a semi-local ring.* R *is quasi-unmixed if and only if the maximal ideals* M_i *in* R *have the same height and* R_{M_i} *is quasi-unmixed* $(i = 1, \cdots, h)$.

Proof. The *only if* part is given by Lemma 2.5, hence assume that the M_i have the same height and that each R_{M_i} is quasi-unmixed. Let M^* be a maximal ideal in the completion R^* of R, and let σ be the natural homomorphism from R^* into $R^*_{M^*}$. Then $I = \text{kernel}\,\sigma = \bigcap_0^\infty M^{*n}$, hence R^*/I

$= R^*_{M^*}$ since R^*/I is a complete local ring. Since R_M is a dense subspace of R^*/I ([3], p. 283), and since R_M is quasi-unmixed, $R^*_{M^*}$ is quasi-unmixed. Therefore, if q is a minimal prime divisor of the zero ideal in R^* which is contained in M^*, then depth $q =$ height $M^* =$ height $M =$ altitude R. Since M^* was arbitrary, R is quasi-unmixed, q. e. d.

LEMMA 2.7. *Let R be a quasi-unmixed semi-local ring, and let $X_1, \cdots, X_n$ be indeterminants. Set $R_n = R[X_1, \cdots, X_n]$. Let $P_1, \cdots, P_r$ be prime ideals in R_n which have the same height, and let S denote the complement in R_n of the union of the P_i. Then R_{nS} is quasi-unmixed.*

Proof. By Lemma 2.6 it may be assumed that $r = 1$. Set $P = P_1$. Since $R_{nS} = R_{nP}$ is a quotient ring of the polynomial ring generated by the X_i over $R_{P \cap R}$, it may be assumed by Lemma 2.5 that R is a quasi-unmixed local ring and that P contains the maximal ideal in R. Let R^* be the completion of R. Set $R^*_n = R^*[X_1, \cdots, X_n]$ and $P^* = PR^*_n$. Since R^*_n/P^* is isomorphic to R_n/P, P^* is a prime ideal in R^*_n. In the proof of (36.8) in ([2], p. 134) it is proved that $R^*_{nP^*}$ and R_{nP} satisfy conditions (1), (2), and (3) of Lemma 2.4. Hence to prove that R_{nP} is quasi-unmixed, it is sufficient by Lemma 2.4 to prove that $R^*_{nP^*}$ is quasi-unmixed. For this, it may be assumed by Lemma 2.5 that P^* is a maximal ideal in R^*_n, hence height $P^* =$ altitude $R^* + n$ (since P^* contains the maximal ideal in R^*, and altitude $R^*_n =$ altitude $R^* + n$ ([2], pp. 27-28)). Let q^{**} be a minimal prime divisor of 0 in $R^{**} = R^*_{nP^*}$, and set $q^* = q^{**} \cap R^*_n$ and $q = q^{**} \cap R^*$. By Lemma 2.2 it is sufficient to prove that depth $q^{**} =$ altitude R^{**} ($=$ altitude $R^* + n$) and that R^{**}/q^{**} is quasi-unmixed. Since the X_i are algebraically independent over R^*, $q^{**} = qR^{**}$ and q is a minimal prime divisor of 0 in R^*. Since R is quasi-unmixed, depth $q =$ altitude R^*. Also R^*_n/q^* is isomorphic to $R^*/q\ [X_1, \cdots, X_n]$. Since the complete local domain R^*/q is a homomorphic image of a regular local ring, R^*_n/q^* is a homomorphic image of a regular domain, hence R^{**}/q^{**} is a homomorphic image of a regular local ring. Therefore R^{**}/q^{**} is unmixed ([2], p. 125). Since R^*_n/q^* is isomorphic to $R^*/q[X]$, height $P^*/q^* =$ altitude $R^*/q + n$ $=$ depth $q + n =$ altitude $R^* + n =$ height P^*. Therefore depth $q^* =$ altitude $R^* + n$, hence depth $q^{**} =$ altitude R^{**}. Therefore R^{**} is quasi-unmixed, hence R_{nS} is quasi-unmixed, q. e. d.

COROLLARY 2.3. *With the same R and R_n of Lemma 2.7, let P and Q be prime ideals in R_n such that $P \subset Q$. Then* height $Q/P =$ height Q $-$ height P.

Proof. By Lemma 2.7 R_{n_0} is quasi-unmixed, hence satisfies the first chain condition for prime ideals (Corollary 2.1). Therefore the conclusion is clear, q. e. d.

PROPOSITION 2.1. *Let R be a quasi-unmixed semi-local ring, let $T=R[a_1,\cdots,a_n]$ be a finitely generated ring over R, let $p_1,\cdots,p_r$ be prime ideals in T which have the same height, and let S denote the complement in T of the union of the p_i. Let A be the kernel of the natural homomorphism from $R_n=R[X_1,\cdots,X_n]$ onto T, and let $P_1,\cdots,P_r$ be prime ideals in R_n which contain A and map modulo A onto $p_1,\cdots,p_r$ respectively. If the minimal prime divisors of A which are contained in $P_1,\cdots,P_r$ have the same height, then T_S is quasi-unmixed.*

Proof. By the condition on the minimal prime divisors of A which are contained in $P_1,\cdots,P_r$, the P_i have the same height (Corollary 2.3). By Lemma 2.7 $R'=R_{n\,(R_n-\overset{r}{\underset{1}{\cup}}P_i)}$ is quasi-unmixed. Again by the condition on the minimal prime divisors of A which are contained in $P_1,\cdots,P_r$ and by Corollary 2.2 $T_S=R'/AR'$ is quasi-unmixed, q. e. d.

COROLLARY 2.4. *With the same R, T, p_i, and S of Proposition 2.1, if T is an integral domain, then T_S is quasi-unmixed.*

Proof. The kernel of the homomorphism from R_n onto T is a prime ideal, hence T_S is quasi-unmixed by Proposition 2.1, q. e. d.

COROLLARY 2.5. *Let R be a quasi-unmixed semi-local integral domain, let T be a finitely generated integral domain over R, and let p be a prime ideal in T. Then T_p is quasi-unmixed.*

Proof. Clear by Proposition2.1.

PROPOSITION 2.2. *Let R be a quasi-unmixed semi-local integral domain. Then R satisfies the altitude formula.*

Proof. By ([3], p. 326) R satisfies the altitude formula if for every $n\geqq 0$ and for every prime ideal p in $R_n=R[X_1,\cdots,X_n]$ (where $X_1,\cdots,X_n$ are algebraically independent over R), R_{np} satisfies the first chain condition for prime ideals. Since R_{np} is quasi-unmixed (Corollary 2.5), R_{np} satisfies the first chain condition for prime ideals (Corollary 2.1), hence R satisfies the altitude formula, q. e. d.

COROLLARY 2.6. *Let R be a quasi-unmixed semi-local ring, let p be a*

prime ideal in R, and let T be an integral domain which is finitely generated over R/p and which is integrally dependent on R/p. Then T is a quasi-unmixed semi-local integral domain.

Proof. Since R/p is quasi-unmixed (Lemma 2.1), by Corollary 2.4 it need only be shown that the heights of the maximal ideals M'_i in T are the same, since $T = T_S$, where S is the complement in T of the union of the maximal ideals in T. Since R/p satisfies the altitude formula (Proposition 2.2), and since T is integrally dependent on R/p, the altitude formula shows that height $M'_i =$ height $(M'_i \cap R/p)$. Since $M'_i \cap R/p$ is a maximal ideal in R/p, and since R/p is quasi-unmixed, the heights of the ideals $M'_i \cap R/p$ are the same, q. e. d.

COROLLARY 2.7. *If R is a quasi-unmixed semi-local ring, then R satisfies the second chain condition for prime ideals.*

Proof. It must be shown that if q is a minimal prime divisor of 0 in R, and if R' is an integral domain which contains R/q and is integrally dependent on R/q, then R' satisfies the first chain condition for prime ideals. For this, it may be assumed that R' is finitely generated over R/q ([2], p. 123). Therefore, by Corollaries 2.6 and 2.1, R satisfies the second chain condition for prime ideals, q. e. d.

COROLLARY 2.8. *Let R be a quasi-unmixed semi-local integral domain, and let T be an affine ring over R. If p is a prime ideal in T, then T_p satisfies the altitude formula and the second chain condition for prime ideals.*

Proof. T_p is quasi-unmixed (Corollary 2.5), hence the conclusions follow from Proposition 2.2 and Corollary 2.7, q. e. d.

COROLLARY 2.9. *If R is an analytically irreducible local domain, and if P is a local quotient ring of an affine ring over R, then P satisfies the altitude formula and the second chain condition for prime ideals.*

Proof. R is quasi-unmixed, q. e. d.

Added in Proof. In a recent personal correspondence Professor Nagata suggested the following proof of the last statement in his proof of Theorem 34.6 in [2] (Corollary 2.6 above).

THEOREM. *If R' is a finite integral extension of a quasi-unmixed semi-local integral domain R, then R' is quasi-unmixed.*

Proof. (1) If R' is contained in the quotient field of R, then the total quotient ring of the completion of R' is the total quotient ring of the completion of R, so the conclusion follows in this case.

(2) The case $R' = R[X]/(f(X))$, where $f(X)$ is monic and irreducible over the quotient field of R. Dropping the assumptions that R is an integral domain and that $f(X)$ is ireducible, we may assume that R is complete. Let p' be a minimal prime division of zero in R', and we have only to show that depth $p' =$ altitude R. Let N be the radical of R. Then N is nilpotent, whence NR' is nilpotent, and $NR' \subseteq p'$. Therefore p' contains pR' for some minimal prime divisor p of zero in R. Now we may assume that $p = 0$. Then no non-zero element of R is a zero-divisor in R' because of the structure of $R' = R[X]/(f(X))$, $f(X)$ being monic. Thus this case is proved.

(3) The general case. There is a sequence of finite integral extensions $R_1, \cdots, R_n$ such that, defining R to be R_0, each R_i $(i = 1, \cdots, n)$ is a finite integral extension of R_{i-1} of type either (1) or (2) above and such that R_n contains R'. R_n is quasi-unmixed by the proof above, whence R' is quasi-unmixed, q. e. d.

UNIVERSITY OF CALIFORNIA, RIVERSIDE.

REFERENCES.

[1] M. Nagata, "On the chain problem of prime ideals," *Nagoya Mathematical Journal*, vol. 10 (1956), pp. 51-64.

[2] ———, *Local Rings*, Interscience Tracts in Pure and Applied Mathematics, 1962.

[3] O. Zariski and P. Samuel, *Commutative Algebra*, Volume II, D. van Nostrand Co., Inc., 1960.

CLASSIFICATION OF NORMAL CONGRUENCE SUBGROUPS OF THE MODULAR GROUP.

By DONALD L. MCQUILLAN.[1]

1. Introduction. Let Γ denote the 2×2 modular group, that is the group of 2×2 rational integral matrices of determinant 1 in which a matrix is identified with its negative. If n is a positive integer then $\Gamma(n)$ denotes the principal congruence subgroup of level n and consists of those elements of Γ congruent modulo n to $\pm I$ where I is the identity matrix. A subgroup G of Γ containing $\Gamma(n)$ is called a congruence subgroup and G is said to be of level n if n is the least such integer. It is the purpose of the present paper to classify all normal congruence subgroups of Γ. Our results will include those of M. Newman [7] who classified certain types of normal congruence subgroups in the case $(n, 6) = 1$. More precisely he showed that if $H(n)$ denotes the normal subgroup of Γ consisting of all elements congruent modulo n to a scalar matrix, and if G is a proper normal congruence subgroup of level n with $(n, 6) = 1$ then $G = H(n)$ whenever G contains $H(n)$.

Since $\Gamma/\Gamma(n) \cong LF(2, n)$ we shall prove our results for this latter group. The case of prime power level is treated in Section 2 and arbitrary level in Section 3.

2. The linear fractional group. The linear fractional group $LF(2, n)$ is defined by $LF(2, n) = SL(2, n)/\pm I$ where $SL(2, n)$ is the special linear group of 2×2 matrices over the ring of integers modulo n and I is the identity matrix. The order of $LF(2, n)$ is $\frac{1}{2}\phi(n)\psi(n)$ where $\phi(n)$ is the Euler function and $\psi(n) = n \prod_{p \mid n} (1 + 1/p)$; the cases $n = 1$ and 2 are exceptional and the orders are then 1 and 6 respectively. It is well known [4] that $T = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ generate $LF(2, n)$ and from this one easily sees that the center of the group consists of all scalar matrices. In particular the center of $LF(2, p^m)$, where p is prime and $m \geqq 0$, reduces

Received March 11, 1964.
Revised December 18, 1964.
[1] Research supported by NSF Grant No. GP-2273.

to the identity except when $p=2$ and $m \geqq 3$ and in this exceptional case the center is a cyclic group of order 2 generted by $\pm \begin{pmatrix} 1+2^{m-1} & 0 \\ 0 & 1+2^{m-1} \end{pmatrix}$. The homomorphism from $LF(2,p^m)$ to $LF(2,p^r)$, $m \geqq r \geqq 0$, defined by reduction modulo p^r will be denoted by f_r^m; this homomorphism is surjective and the kernel will be denoted by K_r^m. We recall ([6], [7]) that K_{m-1}^m is abelian of type (p,p,p) except for the case $p=m=2$ when it is abelian of type $(2,2)$. It is also known [1] that $LF(2,p)$ is simple when $p>3$; $LF(2,3)$ is the alternating group on 4 letters and so it contains a unique normal subgroup V_4 of order 4; $LF(2,2)$ is the symmetric group on 3 letters and thus contains a unique normal subgroup A_3 of order 3. We set $(f_1^m)^{-1}(V_4)=M_m$ and $(f_1^m)^{-1}(A_3)=Q_m$ when $p=3$ and $p=2$ respectively. The following result was proved in [5], [6].

PROPOSITION 1. *The set* $\{K_r^m\}_{r=0}^m$ *gives all normal subgroups of* $LF(2,p^m)$ *when* $p>3$. *When* $p=3$ *there is one other normal subgroup, namely* M_m.

The remainder of this section is devoted to finding the normal subgroups of $LF(2,2^m)$. The center will be denoted by C_m^m $(m \geqq 3)$ and we define $(f_r^m)^{-1}(C_r^r)=C_r^m$ when $m \geqq r \geqq 3$. Since $f_{r-1}^r(C_r^r)=\{1\}$ it is clear that $C_r^m \subset K_{r-1}^m \subset C_{r-1}^m$. When $m \geqq 3$ there is another subgroup of K_{m-1}^m which is normal in $LF(2,2^m)$, namely the group of order 4 consisting of the elements

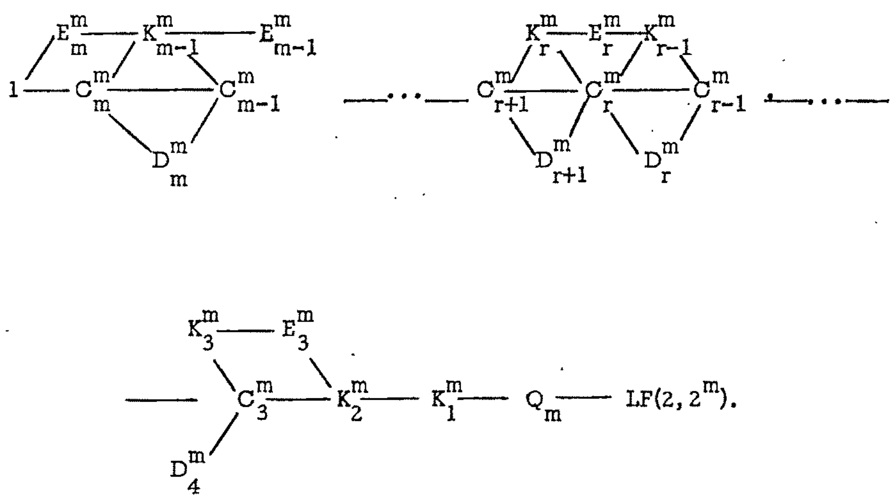
$$\pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm\begin{pmatrix} 1+2^{m-1} & 2^{m-1} \\ 0 & 1+2^{m-1} \end{pmatrix},$$
$$\pm\begin{pmatrix} 1+2^{m-1} & 0 \\ 2^{m-1} & 1+2^{m-1} \end{pmatrix}, \pm\begin{pmatrix} 1 & 2^{m-1} \\ 2^{m-1} & 1 \end{pmatrix}.$$

We denote this group by E_m^m and define $E_r^m=(f_r^m)^{-1}(E_r^r)$ when $m \geqq r \geqq 3$. We note that $E_m^m \cap C_m^m=\{1\}$ and $E_m^m \cdot C_m^m=K_{m-1}^m$. When $m \geqq 4$ the subgroup generated by the element

$$\pm\begin{pmatrix} 1+2^{m-2} & 2^{m-1} \\ 2^{m-1} & 1-2^{m-2} \end{pmatrix}=\pm A$$

say, is denoted by D_m^m; it is easily verified that D_m^m is normal in $LF(2,2^m)$, of order 4, and that $f_{m-1}^m(D_m^m)=C_{m-1}^{m-1}$. In fact $T^{-1}AT=S^{-1}AS=A^3$. Furthermore $D_m^m \cap K_{m-1}^m=C_m^m$ (so that $D_m^m/C_m^m \cong C_{m-1}^{m-1}$ and $D_m^m \cdot K_{m-1}^m=C_{m-1}^m$). If we now define $D_r^m=(f_r^m)^{-1}(D_r^r)$, $m \geqq r \geqq 4$, it is

clear from the preceding remarks and definitions that we have the following diagram of inclusions, intersections and products when $m \geqq 4$.



We note that

$$(C_r^m : K_r^m) = (D_{r+1}^m : C_{r+1}^m) = (K_{r-1}^m : E_r^m) = 2 \quad \text{and}$$
$$(C_r^m : D_{r+1}^m) = (K_r^m : C_{r+1}^m) = (E_r^m : K_r^m) = 4.$$

LEMMA 1. (i) *K_{m-1}^m is contained in the center of K_1^m when $m \geqq 2$;*

(ii) *C_m^m and E_m^m are the only normal subgroups of $LF(2, 2^m)$ properly contained between $\{1\}$ and K_{m-1}^m when $m \geqq 3$.*

Proof. Since K_r^m consists of all elements of $LF(2, 2^m)$ of the form $\pm \begin{pmatrix} 1 + a2^r & b2^r \\ c2^r & 1 + d2^r \end{pmatrix}$ the first statement is immediate. The verification of the second statement is almost trivial when it is recalled that T and S generate $LF(2, 2^m)$.

LEMMA 2. *Let A belong to K_1^m.*

(i) *If $A^2 = 1$ then $f_{m-1}^m(A) = 1$ when $m \geqq 2$;*

(ii) *If $A^2 \in C_m^m$ then $f_{m-1}^m(A) \in C_{m-1}^{m-1}$ when $m \geqq 4$, $f_{m-1}^m(A) = 1$ when $m = 3$;*

(iii) *If $A^2 \in E_m^m$ then $f_{m-1}^m(A) \in E_{m-1}^{m-1}$ when $m \geqq 4$, $f_{m-1}^m(A) = 1$ when $m = 3$.*

Proof. (i) $A \in K_1{}^m$ and $A^2 = 1$ together imply that if $A = \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \epsilon \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ with $a \equiv d \equiv 1$, $b \equiv c \equiv 0 \pmod 2$ and $\epsilon = \pm 1$. It follows that $1 \equiv ad \equiv \epsilon a^2 \pmod 4$ so that $\epsilon = 1$. Then $b \equiv c \equiv 0 \pmod{2^{m-1}}$ and $1 \equiv ad \equiv a^2 \pmod{2^m}$ so that $a \equiv \pm (1 + x2^{m-1}) \pmod{2^m}$ where $x = 0$ or 1. This proves (i). (ii) We may assume here that $A = \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has the property that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \epsilon(1 + 2^{m-1}) \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. The same reasoning used in (i) shows that $\epsilon = 1$ and $b \equiv c \equiv 0 \pmod{2^{m-1}}$. Then $1 \equiv ad \equiv a^2 \pmod{2^m}$ and if $m = 3$ this is impossible. When $m \geqq 4$ we find that $a \equiv \pm (1 + 2^{m-2}) + x2^{m-1} \pmod{2^m}$ where $x = 0$ or 1. This proves (ii). The proof of (iii) is similar.

LEMMA 3. (i) *The order of every element of $K_r{}^m$ is a divisor of 2^{m-r} and $K_r{}^m$ contains an element with precisely this order.*

(ii) *The order of every element of $C_r{}^m$ and $E_r{}^m$ is a divisor of 2^{m-r+1} and each group contains an element with precisely this order.*

(iii) *The order of every element of $D_r{}^m$ is a divisor of 2^{m-r+2} and $D_r{}^m$ contains an element with precisely this order.*

Proof. Since the order of every element $(\neq 1)$ of $K_{m-1}{}^m$ is 2 it follows by induction that the order of an element of $K_r{}^m$ divides 2^{m-r}; similarly for $C_r{}^m$, $E_r{}^m$ and $D_r{}^m$. Furthermore $\pm \begin{pmatrix} 1 & 2^r \\ 0 & 1 \end{pmatrix}$ belongs to $K_r{}^m$ and has order 2^{m-r}. Now choose $A \in C_r{}^m$ so that $f_r{}^m(A)$ has order 2; if $A^{2^{m-r}} = 1$ then by (i) of the previous lemma, $f_{m-1}{}^m(A)^{2^{m-r-1}} = 1$ and continuing thus we conclude that $f_r{}^m(A) = 1$, a contradiction. $E_r{}^m$ and $D_r{}^m$ are treated in the same way.

LEMMA 4. *There is no normal subgroup N of $LF(2, 2^m)$ such that $N \cap K_{m-1}{}^m = \{1\}$ and*

(i) $f_{m-1}{}^m(N) = K_r{}^{m-1}$ *when* $1 \leqq r < m - 1$, *or*

(ii) $f_{m-1}{}^m(N) = C_r{}^{m-1}$ *when* $3 \leqq r \leqq m - 1$, *or*

(iii) $f_{m-1}{}^m(N) = E_r{}^{m-1}$ *when* $3 \leqq r \leqq m - 1$, *or*

(iv) $f_{m-1}{}^m(N) = D_r{}^{m-1}$ *when* $4 \leqq r \leqq m - 1$.

Proof. (i) The hypothesis implies that $N \subset K_r{}^m \subset K_1{}^m$ and $N \cong K_r{}^{m-1}$.

Therefore we have $N\cdot K_{m-1}{}^{m}=K_r{}^{m}$ and this brings a contradiction using Lemas 1 and 3. The proofs of (ii), (iii) and (iv) are exactly the same.

LEMMA 5. *There is no normal subgroup N of $LF(2,2^m)$ such that $f_{m-1}{}^{m}(N)=D_{m-1}{}^{m-1}$ and $N\cap K_{m-1}{}^{m}=C_m{}^{m}$, when $m\geqq 5$.*

Proof. Referring to the definition of $D_m{}^{m}$ one sees that such a normal subgroup must contain an element A of the form

$$A=\pm\begin{pmatrix}1+2^{m-3}+a2^{m-1} & 2^{m-2}+b2^{m-1}\\ 2^{m-2}+c2^{m-1} & 1+2^{m-3}+2^{m-2}+d2^{m-1}\end{pmatrix}.$$

Hence it contains $B=TAT^{-1}$ where $T=\pm\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$. Now

$$B=\pm\begin{pmatrix}1+2^{m-3}+2^{m-2}+(a+c)^{2m-1} & 2^{m-2}+(a+b+c+d)2^{m-1}\\ 2^{m-2}+c2^{m-1} & 1+2^{m-3}+(c+d)2^{m-1}\end{pmatrix}$$

and clearly $f_{m-1}{}^{m}(A^{-1})=f_{m-1}{}^{m}(B)$ so that $AB\in C_m{}^{m}$ i.e. AB is a scalar matrix. This brings an immediate contradiction.

LEMMA 6. *There is no normal subgroup N of $LF(2.2^m)$, $m\geqq 4$, such that $f_{m-1}{}^{m}(N)=E_{m-1}{}^{m-1}$ and $N\cap K_{m-1}{}^{m}=E_m{}^{m}$.*

Proof. Referring to the definition of $E_m{}^{m}$ one sees that such a normal subgroup must contain an element A of the form

$$A=\pm\begin{pmatrix}1+2^{m-2}+a2^{m-1} & 2^{m-2}+b2^{m-1}\\ c2^{m-1} & 1+2^{m-2}+d2^{m-1}\end{pmatrix}$$

where $d\equiv a+1 \pmod 2$. Now $B=TAT^{-1}$ belongs to N, and

$$B=\pm\begin{pmatrix}1+2^{m-2}+(a+c)2^{m-1} & -2^{m-2}+(b+c)2^{m-1}\\ c2^{m-1} & 1+2^{m-2}+(d+c)2^{m-1}\end{pmatrix}$$

Now $f_{m-1}{}^{m}(A)=f_{m-1}{}^{m}(B)$ so that $A^{-1}B\in E_m{}^{m}$. This will bring an immediate contradiction.

LEMMA 7. (i) *if $N\lhd LF(2.2^m)$, $N\cap K_{m-1}{}^{m}=C_m{}^{m}$, and $f_{m-1}{}^{m}(N)=C_{m-1}{}^{m-1}$ then $N=D_m{}^{m}$ when $m\geqq 4$.*

(ii) *There is no normal subgroup N of $LF(2,2^m)$ such that $N\cap K_{m-1}{}^{m}=C_m{}^{m}$ and $f_{m-1}{}^{m}(N)=C_r{}^{m-1}$ or $D_r{}^{m-1}$ when $r\leqq m-2$.*

Proof. $|N|=|D_m{}^{m}|=4$ and both contain $C_m{}^{m}$ so that $N\cap D_m{}^{m}=C_m{}^{m}$ or $D_m{}^{m}$. We show that the former brings a contradiction. Then $N\cdot D_m{}^{m}$ has order 8 and $f_{m-1}{}^{m}(N\cdot D_m{}^{m})=C_{m-1}{}^{m-1}$ so that $N\cdot D_m{}^{m}\cap K_{m-1}{}^{m}$ is a normal

subgroup of $LF(2,2^m)$ of order 4, contained in $K_{m-1}{}^m$ and containing $C_m{}^m$. By Lemma 1 this is impossible.

(ii) If $f_{m-1}{}^m(N) = C_r{}^{m-1}$ then $N \subset C_r{}^m$ and since $N \cap K_{m-1}{}^m = C_m{}^m$ we have $(C_r{}^m : N) = 4$. Referring to the diagram one sees that

$$K_{m-1}{}^m \subset N \cdot D_{r+1}{}^m \subset C_r{}^m$$

since $r \leqq m-2$. Since $f_{m-1}{}^m(N \cdot D_{r+1}{}^m) = C_r{}^{m-1}$ we conclude that $N \cdot D_{r+1}{}^m = C_r{}^m$ and if we set $N_1 = N \cap D_{r+1}{}^m$ then $N_1 \cap K_{m-1}{}^m = C_m{}^m$ and $(D_{r+1}{}^m : N_1) = 4$; furthermore $f_{m-1}{}^m(N_1) = f_{m-1}{}^m(N) \cap D_{r+1}{}^{m-1} = D_{r+1}{}^{m-1}$ since $D_{r+1}{}^m \supset K_{m-1}{}^m$. If $r = m-2$ this gives a contradiction by Lemma 5. Otherwise $K_{m-1}{}^m \subset N_1 \cdot C_{r+1}{}^m \subset D_{r+1}{}^m$ and since $f_{m-1}{}^m(N_1 \cdot C_{r+1}{}^m) = D_{r+1}{}^{m-1}$ we must have $N_1 \cdot C_{r+1}{}^m = D_{r+1}{}^m$. If we set $N_2 = N_1 \cap C_{r+1}{}^m$ then $N_2 \cap K_{m-1}{}^m = C_m{}^m$ and $f_{m-1}{}^m(N_2) = f_{m-1}{}^m(N_1) \cap C_{r+1}{}^{m-1} = C_{r+1}{}^{m-1}$ since $C_{r+1}{}^m \supset K_{m-1}{}^m$. We are therefore back to the original situation with N replaced by N_2 and r by $r+1$. Hence Lemma 5 will eventually bring a contradiction. The fact that $f_{m-1}{}^m(N) = D_r{}^{m-1}$ is impossible is contained in the above.

LEMMA 8. *There is no normal subgroup N of $LF(2,2^m)$ such that $N \cap K_{m-1}{}^m = E_m{}^m$, $m \geqq 3$, and*

(i) $f_{m-1}{}^m(N) = K_r{}^{m-1}$ *when* $1 \leqq r \leqq m-2$, *or*

(ii) $f_{m-1}{}^m(N) = C_r{}^{m-1}$ *when* $3 \leqq r \leqq m-1$, *or*

(iii) $f_{m-1}{}^m(N) = E_r{}^{m-1}$ *when* $3 \leqq r \leqq m-1$, *or*

(iv) $f_{m-1}{}^m(N) = D_r{}^{m-1}$ *when* $4 \leqq r \leqq m-1$.

Proof. (i) If $r = 1$ choose $A \in N$ so that $f_2{}^m(A) \neq 1$. Since $N/E_m{}^m \cong K_1{}^{m-1}$ we have $A^{2^{m-2}} \in E_m{}^m$ by Lemma 3(i) and by repeated applications of Lemma 2(iii) we conclude that $f_2{}^m(A) = 1$, a contradiction. If $r > 1$ choose $A \in N$ so that $f_{r+1}{}^m(A) \notin E_{r+1}{}^{r+1}$ and obtain a contradiction again by Lemma 2(iii).

(ii) Now $N/E_m{}^m \cong C_r{}^{m-1}$ and hence if $A \in N$ then $A^{2^{m-r}} \in E_m{}^m$ by Lemma 3(ii), and applying Lemma 2(iii) we conclude again that $f_r{}^m(A) \in E_r{}^r$. This is a contradiction since $E_r{}^r \cap C_r{}^r = \{1\}$ and we can choose A so that $f_r{}^m(A)$ generates $C_r{}^r$.

(iii) If $r = m-1$ we are finished by Lemma 6. Otherwise $N \subset E_r{}^m$, $(E_r{}^m : N) = 2$ and referring to the diagram one sees at once that $N \cdot K_r{}^m = E_r{}^m$. If $N_1 = N \cap K_r{}^m$ then $N_1 \cap K_{m-1}{}^m = E_m{}^m$ and

$$f_{m-1}{}^m(N_1) = f_{m-1}{}^m(N) \cap K_r{}^{m-1} = K_r{}^{m-1},$$

since $r < m-1$. Part (i) of this lemma gives a contradiction.

(iv) Since $C_r^{m-1} \subset D_r^{m-1}$ the argument used in (ii) applies again.

PROPOSITION 2. *The only normal subgroups of* $LF(2, 2^m)$ *are* $\{K_r^m\}_{r=0}^m$, Q_m, $\{C_r^m\}_{r=3}^m$ $(m \geqq 3)$, $\{E_r^m\}_{r=3}^m$ $(m \geqq 3)$, $\{D_r^m\}_{r=4}^m$ $(m \geqq 4)$.

Proof. The case $m = 2$ can be treated independently and rather easily. We proceed by induction on m. Let N be normal in $LF(2, 2^m)$, $m > 2$, and suppose first that $f_1^m(N) = \{1\}$ so that $N \subset K_1^m$. If $f_{m-1}^m(N) = \{1\}$ then $N = \{1\}$, C_m^m, E_m^m or K_{m-1}^m by Lemma 1(ii). Otherwise, by the induction hypothesis, $f_{m-1}^m(N) = K_r^{m-1}$, $1 \leqq r \leqq m-2$, or C_r^{m-1}, $3 \leqq r \leqq m-1$, or E_r^{m-1}, $3 \leqq r \leqq m-1$, or D_r^{m-1}, $4 \leqq r \leqq m-1$. By Lemmas 4 and 8 $N \cap K_{m-1}^m \neq \{1\}$ or E_m^m. Suppose therefore that $N \cap K_{m-1}^m = C_m^m$. By Lemma 7 we conclude that $N = D_m^m$ or $f_{m-1}^m(N) = K_r^{m-1}$ or E_r^{m-1}. We show that this latter possibility has to be excluded. In the first case $N/C_m^m \cong K_r^{m-1}$, $1 \leqq r \leqq m-2$, and by Lemma 3(i) it follows that $A^{2^{m-r-1}} \in C_m^m$ when $A \in N$. Application of Lemma 2(ii) gives a contradiction. We can exclude $f_{m-1}^m(N) = E_r^{m-1}$ in a similar fashion. There remains only the possibility that $N \cap K_{m-1}^m = K_{m-1}^m$ and then $N = K_r^m$, C_r^m, E_r^m or D_r^m.

Suppose now that $f_1^m(N) = A_3$ so that 3 divides the order of N, $N \subset Q_m$ and, by the induction hypothesis, $f_{m-1}^m(N) = Q_{m-1}$. By comparing orders, it is clear that $N \cdot K_1^m = Q_m$ and hence by one of the isomorphism theorems the order of N is $3s$ where s is the order of $K_1^m \cap N$. It follows from this and from the preceding part of the proof that $K_1^m \cap N \supset K_{m-1}^m$ so that $N = Q_m$. By a similar argument one see that $f_{m-1}^m(N) = LF(2, 2^{m-1})$ implies that $N = LF(2, 2^m)$. The proof of the proposition is complete.

3. The normal subgroups of $LF(2, n)$. We shall in fact determine the normal subgroups of $SL(2, n)$. It will be enough to determine those subgroups which are of level n (cf. Section 1) in the sense of the following

Definition. Let m be a divisor of n and g the natural homomorphism from $SL(2, n)$ to $SL(2, m)$. A subgroup N of $SL(2, n)$ is said to be of level m if there is a subgroup H of $SL(2, m)$ such that $N = g^{-1}(H)$ and m is minimal with this property.

When $p > 3$ it is clear from Proposition 1 of Section 2 that $\{I\}$ and $\{I, -I\}$ are the only normal subgroups of $SL(2, p^m)$ of level p^m; when $p = 3$ and $m = 1$ there is one other, namely the unique normal subgroup of $SL(2, 3)$ which has index 3. This latter group will be denoted by M. When $p = 2$ we shall change slightly the notation of the last section. K_m is

now the kernel of the natural homomorphism from $SL(2, 2^m)$ to $SL(2, 2^{m-1})$ and E_m $(m \geqq 2)$ is the subgroup

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1+2^{m-1} & 2^{m-1} \\ 0 & 1+2^{m-1} \end{pmatrix}, \begin{pmatrix} 1+2^{m-1} & 0 \\ 2^{m-1} & 1+2^{m-1} \end{pmatrix}, \begin{pmatrix} 1 & 2^{m-1} \\ 2^{m-1} & 1 \end{pmatrix},$$

When $m \geqq 4$ the group of order 4 generated by

$$A = \begin{pmatrix} 1+2^{m-2} & 2^{m-1} \\ 2^{m-1} & 1-2^{m-1} \end{pmatrix}$$

is now called D_m and F_m is the group generated by $-A$. In general we shall use $Z(n)$ for the centre of $SL(2, n)$ so that $Z(2) = \{I\}$, $Z(4) = \{I, -I\}$, $Z(2^m) = \{I, -I, B-B\}$ where $m \geqq 3$ and $B = \begin{pmatrix} 1+2^{m-1} & 0 \\ 0 & 1+2^{m-1} \end{pmatrix}$. We set $C_m = \{I, B\}$, $H_m = \{I, -B\}$. Finally Q is the normal subgroup of $SL(2, 2)$ of order 3.

In view of the results of the last section we can state

PROPOSITION 1. *The non-trivial normal subgroups of $SL(2, p^m)$ which are of level p^m are the following*:

$$\begin{array}{lll} Z(p^m), & M & (p=3, m=1) \\ Q\,(p=2, m=1), & & \\ E_m, & \pm E_m & (p=2, m \geqq 2) \\ C_m, & H_m & (p=2, m \geqq 3) \\ D_m, F_m, & \pm D_m & (p=2, m \geqq 4). \end{array}$$

In the statement of the next lemma we use G_m for $SL(2, p^m)$ and $Z(G_m/N)$ for the centre of G_m/N.

LEMMA 1. *When $p=2$ then* $Z(G_m/H_m) = Z(2^m)/H_m \quad (m \geqq 4)$,

$$Z(G_m/N) = \pm K_m/N \qquad (m \geqq 2)$$

if $N = E_m$ *or* $\pm E_m$,

$$Z(G_m/N) = \pm D_m/N \qquad (m \geqq 4)$$

if $N = C_m$, $Z(2^m)$, D_m, F_m, $\pm D_m$.

Proof. $B \bmod N$ belongs to $Z(G_m/N)$ if and only if $BXB^{-1}X^{-1} \in N$ for all X in G_m. On the other hand if B belongs to another normal subgroup N_1 the $BXB^{-1}X^{-1} \in N \cap N_1$ and so if $N \cap N_1 = \{I\}$ than $B \in Z(G_m)$. Since $E_m \cap \pm D_m = \{I\}$ when $m \geqq 4$ it follows easily from the diagram of subgroups in Section 2 that the lemma holds for $N = E_m$ and $\pm D_m$. The other results are now trivial.

Now let G_i be an arbitrary finite group for $i=1,2,\cdots,r$, let $G=\prod_{i=1}^{r} G_i$ and let N be a subgroup of G. We set

$$F_i=\{g_i \mid g_i \in G_i,\ (1,1,\cdots,g_i,\cdots,1) \in N\}$$

and call F_i the i-th foot of N. We set $N_i=\mathrm{pr}_i N$, i.e., the projection of N on G_i.

PROPOSITION 2. *If $G=G_1\times G_2$ then F_i is a normal subgroup of N_i for $i=1,2$ and $N_1/F_1\cong N_2/F_2$. If in addition N is normal in G then F_i is normal in G_i and N_i/F_i is contained in the centre of G_i/F_i.*

Proof. If $n_1 \in N_1$ then there exists $n_2 \in N_2$ such that $(n_1,n_2)\in N$ and so if $f_1\in F_1$ then $(n_1,n_2)^{-1}(f_1,1)(n_1,n_2)$ belongs to N; i.e., $n_1^{-1}f_1n_1\in F_1$. Similarly F_2 is normal in N_2 and if N is normal in G then F_i is normal in G_i, $i=1,2$. Now for each $n_1\in N_1$ we define a subset $\phi(n_1)$ of N_2 by

$$\phi(n_1)=\{n_2 \mid n_2\in N_2,\ (n_1,n_2)\in N\}$$

and for each $n_2\in N_2$ we define a subset $\psi(n_2)$ of N_1 by

$$\psi(n_2)=\{n_1 \mid n_1\in N_1,\ (n_1,n_2)\in N\}.$$

If $n_2\in\phi(n_1)$ one quickly sees that $\phi(n_1)=\phi(1)n_2$ and, since $\phi(1)=F_2$, it follows that $n_1\to\phi(n_1)$ is a map from N_1 to N_2/F_2. That this is a surjective homomorphism follows at once from the definitions, and it is clear that the kernel is precisely F_1. Now if N is normal in G then $\phi(n_1)$ is a normal subset of G_2 in the sense that $g_2\phi(n_1)=\phi(n_1)g_2$ for all g_2 in G_2, since $(1,g_2)^{-1}(n_1,n_2)(1,g_2)\in N$. Factoring by F_2 shows that N_2/F_2 is contained in the centre of G_1/F_1. This completes the proof of the proposition.

Conversely, if F_i is a normal subgroup of G_i, $i=1,2$, and f is an isomorphism from a subgroup N_1^* of $Z(G_1/F_1)$ to a subgroup N_2^* of $Z(G_2/F_2)$ then one can construct a normal subgroup N of $G=G_1\times G_2$ as follows: let N_i be the normal subgroup of G_i which corresponds to N_i^* in G_i/F_i for $i=1,2$, and let ϕ_i be the natural homomorphism from N_i to N_i^*. Then we define N to be the set of all pairs (n_1,n_2) with the property that $f(\phi_1(n_1))=\phi_2(n_2)$. It is clear that N is a normal subgroup of G, that $\mathrm{pr}_i N=N_i$ and that F_i is the i-th foot of N for $i=1,2$. Furthermore the triple (f,F_1,F_2) determine N uniquely and conversely.

LEMMA 2. *If $(n,2)=1$ and F is a subgroup of $Z(n)$ then*

$$Z(SL(2,n)\,/F)=Z(n)/F.$$

Proof. If $A \bmod F$ belongs to the centre of $SL(2,n)/F$ then $AXA^{-1}X^{-1} \in Z(n)$ for all X in $SL(2,n)$ since $F \subset Z(n)$. Therefore $AXA^{-1}X^{-1} \equiv \pm I \pmod{p^s}$ where p is prime and $p^s \parallel n$, and it follows that $A \equiv \pm I \pmod{p^s}$ since the centre of $LF(2,p^s)$ is trivial when $p > 2$. Therefore $A \in Z(n)$ and the lemma follows.

THEOREM 1. *When $(n,6) = 1$ the normal subgroups of level n of $SL(2,n)$ are the subgroups of the centre.*

Proof. Let $n = \prod_{i=1}^{t} p_i^{n_i}$ where p_i are distinct primes; when $t = 1$ the theorem is true and so we proceed by induction on t. Let $m = n/p_1^{n_1}$ and write $SL(2,n) = SL(2,p_1^{n_1}) \times SL(2,m)$. If N is a normal subgroup of level n then $F_1 = \{I\}$ or $\{\pm I\}$. Now the centre of $LF(2,p_1^{n_1})$ is trivial and so by the previous proposition $N_1 = \{I\}$ or $\{\pm I\}$. By the induction hypothesis F_2 is contained in $Z(m)$ and so by Lemma 2 and the same proposition N_2 is also contained in $Z(m)$. Thus N is contained in $Z(p_1^{n_1}) \times Z(m) = Z(n)$ and the theorem is proved.

Now suppose that $n = 3^u \cdot v$ where $(v,6) = 1$, and write $SL(2,n) = SL(2,3^u) \times SL(2,v)$. If N is a normal subgroup of level n then by Proposition 1 $F_1 = \{I\}$, $\{\pm I\}$ or (when $u = 1$) M, and if we exclude the case $F_1 = M$ one shows as above that $N \subset Z(3^u) \times Z(V) = Z(n)$. When $F_1 = M$ then $SL(2,3)/F_1$ is cyclic of order 3 and so $N_1 = F_1$ or $SL(2,3)$; however, by Theorem 1, $F_2 \subset Z(V)$ and so by Lemma 2 N_2/F_2 is a 2-group. Therefore $N_1 = F_1 = M$, $N_2 = F_2$ and $N = M \times F_2$ by Proposition 2.

Definition. A set $A_1, A_2, \cdots, A_w$ of normal subgroups of level n of $SL(2,n)$ will be called a complete set if every normal subgroup of level n is of the form $A_i Z$ where $i = 1, 2, \cdots, w$ and Z is a subgroup of $Z(n)$.

Remark. If $G = \prod_{i=1}^{t} G_i$ is a direct product of groups and A_i is a subgroup of G_i for some i then we shall identify A_i and the subgroup $1 \times 1 \times \cdots \times A_i \times \cdots \times 1 \times 1$ of G.

COROLLARY. *When $(n,2) = 1$ then the set $E = \{I\}$ and M (when $3 \parallel n$) is a complete set of normal subgroups of level n of $SL(2,n)$.*

Finally let $n = 2^m \cdot r$, $(r,2) = 1$, and write

$$SL(2,n) = SL(2,2^m) \times SL(2,r).$$

If $z \in Z(r)$ then $z^2 = I$ and so the element (A,z) and the element $(-A,z)$

each generates a normal subgroup of $SL(2,2^m) \times SL(2,r)$ of order 4. We shall denote these groups by $[A,z]$ and $[-A,z]$. If $z_1, z_2 \in Z(r)$ and $z_1 \neq z_2$ we denote by $[A,z_1,z_2]$ the normal subgroup of $SL(2,2^m) \times SL(2,r)$ consisting of the elements

$$(A,z_1),\ (A^2,z_1),\ (-A,z_2),\ (-A^3,z_2),\ (A^2,I),\ (I,I),\ (-A^2,z_1z_2),\ (-I,z_1z_2).$$

Let now N be a normal subgroup of level n of

$$SL(2,n) = SL(2,2^m) \times SL(2,r).$$

Assume for a moment that if $3 \mid r$ then $9 \mid r$ and that $m \geqq 4$. Then Lemma 2 and the previous corollary show that N_2 is contained in $Z(r)$. If then $F_1 = H_m$ we have by Lemma 1 that N_1 is contained in $Z(2^m)$ and so N is contained in $Z(n)$. If $F_1 = E_m$ then by Lemma 1 again N_1 is contained in $\pm K_m = E_m Z(2^m)$, and so by Proposition 2 and the remarks following it we have:

$$N = F_1 \times F_2$$

or $$N = F_1 \times F_2 \cup F_1\zeta \times F_2 z \quad \text{where}$$

$$\zeta \in Z(2^m) \text{ and } z \in Z(r)$$

or $$N = F_1 \times F_2 \cup F_1\zeta_1 \times F_2 z_1 \cup F_1\zeta_2 \times F_2 z_2 \cup F_1\zeta_3 \times F_2 z_3$$

where $\{I,\zeta_1,\zeta_2,\zeta_3\} = Z(2^m)$ and $\{I,z_1,z_2,z_3\}$ is a subgroup of $Z(r)$ isomorphic to $Z(2^m)$ under the map $\zeta_i \to z_i$. In any case $N = (F_1 \times F_2)Z_0$ where Z_0 is a subgroup of the centre of $SL(2,2^m) \times SL(2,r)$. But $F_2 \subset Z(r)$ and $F_1 = E_m$ so $N = (E_m \times I)Z$ where Z is a subgroup of the centre of $SL(2,2^m) \times SL(2,r)$. A similar argument goes through if $F_1 = \pm E_m$, D_m, $\pm D_m$ or F_m.

Suppose now that $F_1 = Z(2^m)$. Then $N_1 = Z(2^m)$ or $\pm D_m$ by Lemma 1. If $N_1 = Z(2^m)$ then $N \subset Z(n)$ so assume $N_1 = \pm D_m = Z(2^m) \cup Z(2^m)A$. By Proposition 2 we must have $N_2 = F_2 \cup F_2 z$ where $z \in Z(r)$ and

$$\begin{aligned} N &= Z(2^m) \times F_2 \cup Z(2^m)A \times F_2 z \\ &= (Z(2^m) \times F_2)[A,z] \\ &= [A,z]Z \text{ where } Z \text{ is contained in the centre.} \end{aligned}$$

If $F_1 = C_m$ then $N_1 = C_m$, $Z(2^m)$, D_m, F_m or $\pm D_m$, arguments similar to the above show that N is contained in the centre or $N = [A,z]Z$, $[-A,z]Z$ or $[A,z_1,z_2]Z$.

The cases $m = 2$ and $m = 3$ are treated similarly. When $m = 1$ there

is one other possibility to be considered, namely $F_1 = Q$. Then $N_1 = Q$ or $SL(2,2)$ so that $N = Q \times F_2$ or $N = Q \times F_2 \cup QT \times F_2 z$ where $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ in $SL(2,2)$ and $z \in Z(r)$. In the latter case we have $N = (Q \times F_2)[T,z]$ where $[T,z]$ is the cyclic (non-normal) subgroup of $SL(2,2) \times SL(2,r)$ generated by the element (T,z). Therefore, $N = (Q \times I)[T,z]Z$. We shall denote the normal subgroup $(Q \times I)[T,z]$ of $SL(2,2) \times SL(2,r)$ by $[Q,T,z]$.

There remains the possibility that $3 \parallel r$ and $F_2 = M$. But then the same arguments show that $N = N_1 M$, where N_1 is one of the normal subgroups already obtained. We can state the

MAIN THEOREM. *The following is a complete set of normal subgroups of level n of $SL(2,n)$:*

$\{I\}$

Q, $[Q,T,z]$ *when* $2 \parallel m$

E_m when $4 \mid m$

F_m, D_m, $[A,z]$, $[-A,z]$, $[A,z_1,z_2]$ *when* $2^4 \mid n$ *and, if* $3 \parallel n$, *the above groups multiplied by* M.

MADISON, WISCONSIN.

REFERENCES.

[1] L. E. Dickson, ***Linear Groups***, New York, 1958.
[2] G. Gierster, "Die Untergruppen der Galois'schen Gruppe der Modulargleichungen für den Fall eines primzahligen Transformationsgrades," ***Mathematische Annalen***, Bd. 18 (1881), pp. 319-365.
[3] ———, "Über die Galois'sche Gruppe Modulargleichungen, wenn der Transformationsgrad Potenz einer Primzahl > 2 ist," *Mathematische Annalen*, Bd. 26 (1886), pp. 309-368.
[4] R. C. Gunning, *Lectures on Modular Forms*, Princeton, 1962.
[5] D. McQuillan, "Some structure theorems for $SL(2,n)$," Abstracts, *International Congress*, Stockholm, 1962.
[6] ———, "Some results on the linear fractional group," *Illinois Journal of Mathematics* (to appear).
[7] M. Newman, "Normal congruence subgroups of the modular group," *American Journal of Mathematics*, vol. 85 (1963), pp. 419-427.

ON THE HOMOTOPY GROUPS OF THE WEDGE OF SPHERES.

By GERALD J. PORTER.

In a paper with a similar title [4] Hilton proved that $\pi_p(S^{n_1} \vee \cdots \vee S^{n_k})$ is isomorphic to a direct sum of homotopy groups of spheres. We show that a similar result is true for the 'fat wedge' of spheres.

Corresponding to each element of $\pi_*(S^{n_1} \vee \cdots \vee S^{n_k})$ there is a primary homotopy operation on k variables. Similarly each element of the homotopy of the fat wedge of k-spheres corresponds to a $(k-1)$-ary homotopy operation on k variables. Let $T_i = T_i(S^{n_1}, \cdots, S^{n_k})$ be the subset of $S^{n_1} \times \cdots \times S^{n_k}$ consisting of those points with at *least* i coordinates at base points. T_0 is the cartesian product, T_{k-1} is the union of spheres studied by Hilton, and T_1 is the fat wedge which is studied in this paper. It will always be assumed that $n_i > 1$ for each i and $k > 2$.

Hilton's calculation of $\pi_*(T_{k-1})$ was done by relating $\pi_*(T_{k-1})$ to the Pontryagin ring $H_*(\Omega T_{k-1})$. This relation holds for more general spaces than T_{k-1}. In particular if X is a simply connected CW-complex we give conditions on the Pontryagin ring $H_*(\Omega X)$ which ensure that $\pi_p(X)$ is isomorphic to a direct sum of homotopy groups of spheres for each p. The proof of this relation is essentially the same as the proof of Hilton's theorem. The relation is stated and proved in Section 1.

To apply the result of Section 1 to T_1 it is necessary to know $H_*(\Omega T_1)$. This is calculated in Section 3 using the Serre spectral sequence.

Assuming the result of Section 3, we show in Section 2 that $H_*(\Omega T_1)$ satisfies the hypothesis of the theorem proved in Section 1. To do this we give a free additive basis of commutators for the ring

$$Z[x_1, \cdots, x_n] * Z[y_1] * \cdots * Z[y_k],$$

where $*$ stands for the free product of rings.

The calculation of $H_*(\Omega T_1)$ was contained in the author's doctoral dissertation written at Cornell University under the direction of Professor William Browder. The author wishes to express his appreciation to Professor Browder for his guidance.

The author has shown, by methods different from those used in this

Received May 18, 1964.

paper, that the calculation of $\pi_*(T_1)$ can be extended to the fat wedge of arbitrary suspensions to get a result similar to Milnor's extension of Hilton's theorem [5].

Section 1. A relation between $\pi_*(X)$ and $H_*(\Omega X)$. For a graded ring R, define

$$[a,b]=(-1)^p(a\cdot b-(-1)^{pq}b\cdot a)$$

for elements a and b of gradation p and q respectively.

Let R be freely generated by elements $e_1,\cdots,e_n$. A set of R-basic products is defined as follows. The R-basic products of weight 1 are $e_1,\cdots,e_n$. These are ordered by setting $e_1<\cdots<e_n$. Assume the R-basic products of weight less than w have been defined and are ordered. An R-basic product of weight w is a bracket $[a,b]$ where a and b are R-basic products of weight u and v respectively, $u+v=w$, $a<b$, and if b is defined by the bracket $[c,d]$, $c\leqq a$. The products of weight w are then ordered arbitrarily among themselves and are greater than products of lesser weight. This then defines a set of R-basic products.

We note that the ordering of the basic products of a given weight is arbitrary. For another given ordering we would have a different set of R-basic products.

Let $b_1,\cdots,b_k$ be elements of a set of R-basic products. A monomial $b_1\cdots b_k$ is said to be ordered if $i<j$ implies $b_i\leqq b_j$. Hilton [4] has shown that the Witt-Magnus theorem is valid for a general class of bracket operations of which the bracket operation given above is a member. In particular:

THEOREM 1.1. *If $e_1,\cdots,e_n$ are free generators of the free associative ring R, then the ordered monomials in a set of R-basic products form a free additive basis for R.*

Let S be a finitely generated graded ring. We say that S is *basic* if there exists a freely generated graded ring R and a set of R-basic products such that

1. $S\approx R/I$ for some graded ideal I.

2. There exists a subset T of the R-basic products which satisfies the following property: If $\rho(T)$ is the image of T in S, then the ordered monomials in elements of $\rho(T)$ form a free additive basis for S.

We call $\rho(T)$ a set of S-basic products. If S is a freely generated ring, Theorem 1.1 ensures that the two definitions of S-basic products agree.

We note that the choice of a set of S-basic products may not be unique. In any case, it follows from condition 2 that the number of S-basic products of gradation q does not depend upon the choice of a set of S-basic products.

We shall assume throughout that either $R_0 = 0$ or $R_0 \approx Z$ depending upon whether or not R has a unit. In the latter case we say that R is basic if $\sum_{q>0} R_q$ is basic.

It is shown in Section 2 that any ring of the form

$$Z[x_1, \cdots, x_n] * Z[y_1] * \cdots * Z[y_k]$$

is basic, where $R * S$ is the free product of R and S. On the other hand it is easily seen that the ring $Z[x]/x^n$ is not basic for any $n > 1$.

For a simply connected CW-complex, X, let τ be the composite map

$$\pi_j(X) \to \pi_{j-1}(\Omega X) \xrightarrow{h} H_{j-1}(\Omega X).$$

where the first map is the canonical isomorphism and h is the Hurewicz homomorphism.

We say that X is *basic* if $H_*(\Omega X)$ is basic and there is a set of $H_*(\Omega X)$-basic products which is contained in the image of τ.

For each $H_*(\Omega X)$-basic product, $x \in H_{n-1}(\Omega X)$, choose $b \in \pi_n(X)$ such that $\tau(b) = x$. We call b an X-basic product and assign to it the same weight as x. The set of such b's is called a set of X-basic products and is in 1-1 correspondence with the set of $H_*(\Omega X)$-basic products.

Let [,] stand for the Whitehead product in $\pi_*(X)$ and also for the operation defined above in the graded Pontryagin ring $H_*(\Omega X)$. Samelson [7] has shown that

$$\tau[a, b] = [\tau a, \tau b] \text{ for } a, b \in \pi_*(X).$$

Remark. Assume that $H_*(\Omega X)$-basic products satisfy the following property: Whenever $[x, y]$ is an $H_*(\Omega X)$-basic product, x and y are also $H_*(\Omega X)$-basic products. Then if each $H_*(\Omega X)$-basic product of weight one is in the image of τ, the Samelson theorem (see above) ensures that each $H_*(\Omega X)$-basic product is in the image of τ. Moreover, once a set of X-basic products of weight 1 have been chosen, the X-basic products of higher weight may be chosen using the Whitehead product.

For the remainder of this section we will assume that a set of X-basic products has been chosen. For each basic product, $a \in \pi_{n_a}(X)$, a homomorphism $f_a \colon \pi_p(S^{n_a}) \to \pi_p(X)$ is given by setting $f_a(\beta) = a \cdot \beta$. Let $f \colon \sum_a \pi_p(S^{n_a}) \to \pi_p(X)$ be defined by $f = \sum_a f_a$, where the sum is taken over the set of X-basic products.

THEOREM 1.2. *If X is basic, the homomorphism $f\colon \sum_a \pi_p(S^{n_a}) \to \pi_p(X)$ is an isomorphism for each integer p.*

Theorem 1.2 was proven by Hilton [4] in the case $X = S^{n_1} \vee \cdots \vee S^{n_k}$. It is easily seen that his proof generalizes to the case in which X is basic. We sketch the proof below.

We first remark that since X is basic, $H_*(\Omega X)$ is finitely generated as a ring and $\sum_a \pi_p(S^{n_a})$ is a finite direct sum for each p.

Proof of Theorem 1.2. Let $\Omega_a = \Omega S^{n_a}$ for each X-basic product a. $H_*(\Omega S^{n_a})$ is a polynomial algebra on one generator y_a of dimension $n_a - 1$. Let L be the direct limit of finite ordered products of the Ω_a corresponding to the set of X-basic products. $H_*(L)$ is easily seen to have a basis consisting of elements of the form $y_{a_1}{}^{m_1} \otimes \cdots \otimes y_{a_i}{}^{m_i}$.

Each X-basic product defines a multiplicative map $g_a\colon \Omega_a \to \Omega X$ in the obvious way. Let x_a be the image of y_a under $(g_a)_*$. It is easily seen that $x_a = \tau a$. Hence the set of x_{a_i} is simply the set of $H_*(\Omega X)$-basic products.

Since g_a is multiplicative, $(g_a)_*(y_a{}^m) = x_a{}^m$ and $g_*(y_{a_1}{}^{m_1} \otimes \cdots \otimes y_{a_i}{}^{m_i})$ is the *ordered monomial* $x_{a_1}{}^{m_1} \cdots x_{a_i}{}^{m_i}$. The set of such monomials forms a basis for $H_*(\Omega X)$ since $H_*(\Omega X)$ is basic and the x_{a_i} are the $H_*(\Omega X)$-basic products. Thus $g_*\colon H_*(L) \to H_*(\Omega X)$ is an isomorphism. If ΩX were simply connected we could apply the Whitehead theorem. Hilton has shown that even if ΩX is not simply connected, g still induces an isomorphism between $\pi_{p-1}(L)$ and $\pi_{p-1}(\Omega X)$ for each p.

Since $\pi_{p-1}(L) \approx \sum_a \pi_p(S^{n_a})$ and $\pi_{p-1}(\Omega X) \approx \pi_p(X)$ it follows from the definitions of g and f that

$$f\colon \sum_a \pi_p(S^{n_a}) \to \pi_p(X)$$

is an isomorphism.

COROLLARY. *If X is basic, $H_*(\Omega X)$ is isomorphic to the universal enveloping algebra of $h\pi_*(\Omega X)$, where h is the Hurewicz homomorphism.*

Proof. We first remark that $h\pi_*(\Omega X)$ is a graded sub-Lie algebra of $H_*(\Omega X)$ and therefore its universal enveloping algebra is defined.

Let $\mathfrak{g} = h\pi_*(\Omega X)$ and $\mathfrak{g}_a = h\pi_*(\Omega S^{n_a})$ for each X-basic product a. Theorem 1.2 implies that $\mathfrak{g} \approx \sum_a \mathfrak{g}_a$. Thus $U(\mathfrak{g})$, the universal enveloping algebra of $\mathfrak{g}$, is isomorphic to $\otimes_a U(\mathfrak{g}_a)$, where $U(\mathfrak{g}_a)$ is the universal enveloping algebra of $\mathfrak{g}_a$.

The Corollary follows by noting that $U(\mathfrak{g}_a)$ is isomorphic to $H_*(\Omega S^{n_a})$ for each basic product a.

Section 2. Calculation of $\pi_*(T_1(S^{n_1},\cdots,S^{n_k}))$. Let

$$T_i = T_i(S^{n_1},\cdots,S^{n_k})$$

be the subset of $S^{n_1}\times\cdots\times S^{n_k}$ consisting of those points with at *least* i coordinates at base points. It is always assumed that each $n_i > 1$ and $k > 2$.

To apply Theorem 1.2 we must show that T_1 is basic. For any two rings R and S we denote the free product of R and S by $R * S$. The following theorem is proved in Section 3.

THEOREM 2.1. *$H_*(\Omega T_1) \approx H_*(\Omega T_0) * H_*(\Omega S^{N-1})$ where $N = \sum_{i=1}^{k} n_i$ and the isomorphism is a ring isomorphism.*

We recall that $H_*(\Omega T_0) \approx Z[\alpha_1,\cdots,\alpha_k]$ and $H_*(\Omega S^{N-1}) \approx Z[\beta]$ where $\dim \alpha_i = n_i - 1$ and $\dim \beta = N - 2$. The fact that $H_*(\Omega T_1)$ is basic follows from the following theorem.

THEOREM 2.2. *$Z[x_1',\cdots,x_n'] * Z[y_1'] * \cdots * Z[y_m']$ is basic.*

Proof. Let R be the free associative ring generated by elements $x_1,\cdots,x_n, y_1,\cdots,y_m$. Assume R is graded so that gradation x_i = gradation x_i' and gradation $y_j = y_j'$ for $1 \leqq i \leqq n$ and $1 \leqq j \leqq m$. Let $\mathfrak{R}$ be a set of R-basic products with $x_1 < \cdots < x_n < y_1 < \cdots < y_m$.

Define $B \subset \mathfrak{R}$ by $B = \{[x_i, x_j], i < j\}$, and let I be the graded two sided ideal of R generated by B. It is clear that for elements of positive gradation

$$R/I \approx Z[x_1',\cdots,x_n'] * Z[y_1'] * \cdots * Z[y_m']$$

where x_i' is the image of x_i in R/I and y_j' is the image of y_j.

Let $S \subset \mathfrak{R}$ be the set of elements of $\mathfrak{R}$ which are neither in B nor defined using an element of B. Denote the image of S in R/I by $\mathcal{S}$ and let $\mathfrak{I} = \mathfrak{R} - S$. We shall show that $\mathcal{S}$ is a set of R/I-basic products. A monomial in the elements of $\mathfrak{R}$ is called an $\mathfrak{I}$-monomial if some factor is in $\mathfrak{I}$. Since all the $\mathfrak{I}$ monomials are elements of I, it is clear that the ordered monomials in elements of $\mathcal{S}$ are a set of additive generators for R/I. The independence of the ordered monomials in the elements of $\mathcal{S}$ is a consequence of the following lemma.

LEMMA 2.3. *The set of ordered $\mathfrak{I}$-monomials forms an additive basis for I.*

We defer the proof of Lemma 2.3 and give the application of Theorem 2.2. It follows from Theorem 2.2 that $H_*(\Omega T_1)$ is basic. Thus to show

that T_1 is basic it suffices by the remark on p. 299 to prove the following lemma.

LEMMA 2.4. *There are elements $i_1, \cdots, i_k$ and W in $\pi_*(T_1)$ such that $\tau(i_j) = \alpha_j$ and $\tau(W) = \beta$ where $H_*(\Omega T_1) \approx Z[\alpha_1, \cdots, \alpha_k] * Z[\beta]$.*

Proof of Lemma 2.4. It suffices to find homotopy elements $i_1, \cdots, i_k$ and W such that $\tau(i_j) = \pm\alpha_j$ and $\tau(W) = \pm\beta$. Let $i_j\colon S^{n_j} \to T_1$ be the canonical injection. It is easily seen that $\tau(i_j) = \pm\alpha_j$. It remains to find W.

The higher order Whitehead product was defined and studied in [6]. We recall the basic definition. Let ω be a generator of $\pi_N(T_0, T_1)$, which is isomorphic to the integers by the relative Hurewicz theorem.

For each map $g\colon T_1 \to X$, the k-th order Whitehead product, $W(g) \in \pi_{N-1}(X)$, is defined to be $g_*\partial\omega$, where ∂ is the boundary operator in the homotopy sequence of the pair (T_0, T_1).

Let $i\colon T_1 \to T_1$ be the identity. Let $W = W(i) = \partial\omega$. We show that $\tau(W) = \pm\beta$.

Consider the following diagram

$$\begin{tikzcd}
\pi_N(T_0, T_1) \arrow[r, "\partial"] \arrow[d, "A'"] & \pi_{N-1}(T_1) \arrow[d, "A"] & \\
\pi_{N-1}(\Omega T_0, \Omega T_1) \arrow[r, "\partial'"] \arrow[d, "h'"] & \pi_{N-2}(\Omega T_1) \arrow[d, "h"] & \\
H_{N-1}(\Omega T_0, \Omega T_1) \arrow[r, "d"] & H_{N-2}(\Omega T_1) \arrow[r, "j_*"] & H_{N-2}(\Omega T_0)
\end{tikzcd}$$

where the rows are segments of exact homotopy and homology sequences, A and A' are the canonical maps and h and h' are Hurewicz homomorphisms.

Samelson [7] has shown that the above diagram commutes and that A' is an isomorphism for all j. It follows that $\pi_j(\Omega T_0, \Omega T_1) = 0$ for $j < N-1$. Since the pair $(\Omega T_0, \Omega T_1)$ is $(N-1)$-simple, h' is an isomorphism by the relative Hurewicz theorem.

Using the exactness of the lower row and the fact that $h'A'$ is an isomorphism we see that $dh'A'\omega$ generates the kernel of j_*. However it follows from Theorem 2.1 that β generates the kernel of j_*. Therefore $\pm\beta = dh'A'\omega = hA\partial\omega = \tau(W)$.

COROLLARY (to the proof). *If $g\colon T_1 \to X$, $\tau(W(g)) = \pm g_*\beta$.*

Thus T_1 is basic and by Theorem 1.2 we have

THEOREM 2.5. *Given a set of T_1-basic products they induce an isomorphism between $\sum_{a} \pi_p(S^{n_a})$ and $\pi_p(T_1)$, where the sum is taken over the given set of* T_1-*basic products.*

Using Theorem 2.2 we may describe the T_1-basic products as follows.

$$\begin{aligned}
&\text{weight } 1\text{—}i_1, i_2, \cdots, i_k, W(i).\\
&\text{weight } 2\text{—}[i_1, W(i)], \cdots, [i_k, W(i)].\\
&\text{weight } 3\text{—}[i_j, [i_l, W(i)]] \quad j \geqq l \text{ and}\\
&\qquad\qquad [W(i), [i_j, W(i)]] \quad j = 1, \cdots, k.
\end{aligned}$$

The products of weight greater than 2 may be defined inductively in a manner similar to the original definition of R-basic products in Section 1.

Theorem 2.5 implies, for example, that if $p < (k+2)n - 3$

$$\pi_p(T_1(S^{n_1}, \cdots, S^{n_k})) \approx \sum_{j=1}^{k} \pi_p(S^n) \oplus \pi_p(S^{kn-1}) \oplus \sum_{j=1}^{k} \pi_p(S^{(k+1)n-2}).$$

To complete the proof of Theorem 2.5 we must prove Lemma 2.3. This is done using the collection process of P. Hall.

Proof of Lemma 2.3. To prove the lemma it suffices to show that each $\mathfrak{A}$-monomial may be written as the sum of ordered $\mathfrak{A}$-monomials. Let $b_i \in \mathfrak{R}$, $1 \leqq i \leqq n$. We define the degree of $b_1 \cdots b_n$ to be n and its disorder to be the number of pairs (i, j), with $1 \leqq i < j \leqq n$, such that $b_i > b_j$. We prove the above assertion by double induction over the degree and the disorder. If either the degree is one or the disorder is zero the lemma is true by definition. Assume, inductively, that the assertion is proven for all $\mathfrak{A}$-monomials of degree less than n and all $\mathfrak{A}$-monomials of degree n and disorder less than r. Let $b_1 \cdots b_n$ be an $\mathfrak{A}$-monomial of degree n and disorder r.

Let b_{i+1} be the first occurrence of the smallest factor in disorder. Thus $b_i > b_{i+1}$. It follows from the definition of [,] that there are integers λ_1 and λ_2, $|\lambda_i| = 1$, such that

$$b_i b_{i+1} = \lambda_1 b_{i+1} b_i + \lambda_2 [b_{i+1}, b_i]. \tag{2.3.1}$$

Thus

$$\begin{aligned}
b_1 \cdots b_n = {} & \lambda_1 b_1 \cdots b_{i-1} b_{i+1} b_i b_{i+2} \cdots b_n \\
& + \lambda_2 b_1 \cdots b_{i-1} [b_{i+1}, b_i] b_{i+2} \cdots b_n.
\end{aligned}$$

Assume that if $b_i \in \mathfrak{A}$, $[b_{i+1}, b_i]$ is the sum of $\mathfrak{A}$-monomials of degree 1. We then replace $[b_{i+1}, b_i]$ by this sum in the above equation. We may assume that $b_1 \cdots b_n$ arose from a monomial $a_1 \cdots a_l$, where each a_i is either a

generator of R or an element of B, by repeated applications of the above process. We must show that if $b_i \notin \mathfrak{I}$ then $[b_{i+1}, b_i] \in \mathcal{R}$. If b_i is a basic product of weight 1 then $[b_{i+1}, b_i] \in \mathcal{R}$ since $b_{i+1} < b_i$. Assume $b_i = [c, d]$. Since $b_i \notin \mathfrak{I}$, $[c, d]$ arose from an earlier application of the collection process. Therefore $c \leqq b_{i+1}$ and $[b_{i+1}, b_i] \in \mathcal{R}$. Furthermore if $b_{i+1} \in \mathfrak{I}$ so is $[b_{i+1}, b_i]$. Thus $b_1 \cdots b_n$ is the sum of $\mathfrak{I}$-monomials with either degree $(n-1)$ or degree n and disorder $(r-1)$. By induction $b_1 \cdots b_n$ can then be written as the sum of ordered $\mathfrak{I}$-monomials.

It remains to be shown that if $b_i \in \mathfrak{I}$ then $[b_{i+1}, b_i]$ is the sum of $\mathfrak{I}$-monomials of degree 1. We remark that up to this point we have not used the definition of B and have used the definition of $[\ ,\ ]$ only for equation (2.3.1).

Let $b_i \in \mathfrak{I}$ be equal to $[c, d]$. If $c \leqq b_{i+1}$, $[b_{i+1}, b_i] \in \mathcal{R}$ and we are done. Assume $c > b_{i+1}$. For our particular definition of $[\ ,\ ]$ there exists a Jacobi identity, i.e. there are integers ϵ_1, ϵ_2, and ϵ_3 depending upon b_i and b_{i+1} such that

$$(2.3.2) \qquad [b_{i+1}, b_i] = \begin{cases} \epsilon_1[c, [b_{i+1} d]] + \epsilon_2[d, [b_{i+1}, c]] & \text{if } d < [b_{i+1}, c] \\ \epsilon_1[c, [b_{i+1}, d]] + \epsilon_3[[b_{i+1}, c], d] & \text{if } d > [b_{i+1}, c]. \end{cases}$$

In particular $|\epsilon_j| = 1$ for $j = 1, 2, 3$.

If either $c \in \mathfrak{I}$ or $d \in \mathfrak{I}$ and $d \neq [b_{i+1}, c]$ then (2.3.2) shows that $[b_{i+1}, b_i]$ is a sum of $\mathfrak{I}$-monomials of degree 1. If $c \notin \mathfrak{I}$ and $d \notin \mathfrak{I}$, then $b_i \in B$. Therefore $[c, d] = [x_i, x_j]$ for $i < j$. Since $b_{i+1} < c$, $b_{i+1} = x_l$ for $l < i$. By the definition of B both $[b_{i+1}, c]$ and $[b_{i+1}, d]$ are also in B and we are done. The only case remaining is $[b_{i+1}, c] = d$, i.e. $b_i = [c, [b_{i+1}, c]]$. We show that no such element can arise. There are three ways in which b_i could have arisen i) $b_i \in B$, ii) b_i came from an earlier collection involving c, or iii) b_i came via the Jacobi identity from a collection of $[c, c]$. These are all impossible. i) is impossible by the definition of B; ii), since $c > b_{i+1}$; and iii), since $[c, c] \notin \mathcal{R}$. This completes the proof of the lemma.

Actually we have proven a slightly more general result. Let γ_1 and γ_2 be two mappings of $R \times R$ into the set ± 1. We define the quasi-commutator of a and b by

$$[a, b] = \gamma_1(a, b) a \cdot b - \gamma_2(a, b) b \cdot a.$$

If there exists mappings ϵ_1, ϵ_2, and ϵ_3 of $R \times R \times R$ into the set ± 1 such that for all $(a, b, c) \in R \times R \times R$

$$\epsilon_1[a[b, c]] + \epsilon_2[b, [a, c]] + \epsilon_3[c, [a, b]] = 0$$

we call $[\ ,\ ]$ a J-quasi-commutator.

We may then define $\mathfrak{R}$, a set of R-basic products, in terms of the quasi-commutator. The explicit nature of our quasi-commutator was used in the above proof at two steps, formulas (2.3.1) and (2.3.2). It is easily seen that all we needed at these steps was a J-quasi-commutator.

The definition of the subset, B, of $\mathfrak{R}$ was used twice in the proof of Lemma 2.3, the critical properties that B satisfied were

1. If $[a,b]\in B$ and $c\in\mathfrak{R}$, $c<a$, then $[a,[c,b]]\in B$ and either $[[c,a],b]\in B$ or $[b,[c,a]]\in B$.

2. There are no elements of the form $[u,[v,u]]\in B$.

It is easily checked that if 2 were not true the following condition would have sufficed.

2′. $[a,a]=0$ for all a.

We call a subset B of $\mathfrak{R}$ closed if either properties 1 and 2 or 1 and 2′ hold.

Let $\mathcal{S}$ be defined in terms of $\mathfrak{R}$ and B as in Theorem 2.2. We have shown

THEOREM 2.6. *Let $\mathfrak{R}$ be a set of R-basic products defined by a J-quasi-commutator. Let B be a closed subset of $\mathfrak{R}$. Then $\mathcal{S}$ is a set of R/I-basic products where I is the two sided ideal generated by B.*

Question. Is there a classification theorem for basic rings?

Section 3. Calculation of $H_*(\Omega T_1(S^{n_1},\cdots,S^{n_k})$. In this section we prove Theorem 2.1. The proof uses the Serre spectral sequence [8] and the comparison theorem of Zeeman [9] as modified by Dyer and Lashof [2].

THEOREM 3.1. *Suppose $\{h^r\}:\{'E^r\}\to\{E^r\}$ is a homomorphism of homology spectral sequences and*

$$'\chi_{a,b}:{}'E^2_{a,b}\to{}'E^2_{a,0}\otimes{}'E^2_{0,b}\ \textit{and}$$
$$\chi_{a,b}:E^2_{a,b}\to E^2_{a,0}\otimes E^2_{0,b}$$

are isomorphisms such that

$$\chi_{a,b}h^2_{a,b}=(h^2_{a,0}\otimes h^2_{0,b})'\chi_{a,b}.$$

If $h^2_{a,0}$ is an isomorphism for $0\leqq a\leqq n$ and $h^\infty_{a,b}$ is an isomorphism for $0\leqq a+b\leqq n$ and $b\leqq n-2$ then

i) $h^2_{0,b}$ is an isomorphism for $0 \leqq b \leqq n-2$,

ii) $h^r_{a,b}$ is an epimorphism if either $a+b \leqq n+1$, $a \leqq r-2$ and $b \leqq n-2$ or $b \leqq n-r$ and $a \leqq n$,

iii) $h^r_{a,b}$ is a monomorphism if $b \leqq n-2$ and either $a+b \leqq n-a$ and $n-(r-1) \leqq a$ or $a \leqq n-(r-1)$.

Let PX be the subset of X^I consisting of those paths f such that $f(0) = *$, the base point in X. Define $\pi\colon PX \to X$ by setting $\pi(f) = f(1)$. (PX, π, X) is then a fibre space with fibre ΩX. Let $T_1 = T_1(S^{n_1}, \cdots, S^{n_k})$ for $n_i > 1$, $1 \leqq i \leqq k$, $k > 2$, and $n_1 \leqq n_2 \leqq \cdots \leqq n_k$. Let $\{A^r\}$ be the spectral sequence associated with (PT_1, π, T_1). Since T_1 is simply connected,

$$A^2_{p,q} \approx A^2_{p,0} \otimes A^2_{0,q} \approx H_p(T_1) \otimes H_q(\Omega T_1).$$

To apply Theorem 3.1, there are two principal steps. First, a model spectral sequence having the desired $\{E^2\}$ and $\{E^\infty\}$ terms must be defined. Secondly, a homomorphism of spectral sequences, $\{h^r\}\colon \{E^r\} \to \{A^r\}$, must be defined such that $h^2_{p,0}$ is an isomorphism for all p, and h^∞ is an isomorphism. Theorem 3.1 would then imply

$$(\sum_q h^2_{0,q}) : \sum_q E^2_{0,q} \approx \sum_q A^2_{0,q} \approx H_*(\Omega T_1).$$

Construction of the model spectral sequence.

THEOREM 3.2. *There exists a spectral sequence $\{E^r\}$, $r \geqq 2$ satisfying*

1. $\sum_{q \geqq 0} E^2_{0,q} \approx Z[\alpha_1, \cdots, \alpha_n] * Z[\beta]$ *as rings*
2. $E^2_{p,0} \approx H_p(T_1)$ *for all* p
3. $E^2_{p,q} \approx E^2_{p,0} \otimes E^2_{0,q}$
4. $E^\infty_{p,q} = 0$ *if* $(p,q) \neq (0,0)$.

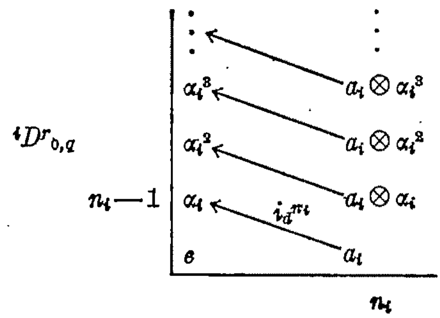
For any spectral sequence, $\{I^r\}$, let

$$\bar{I}^r_{p,q} = \begin{cases} I^r_{p,q} & \text{if } (p,q) \neq (0,0) \\ 0 & \text{if } (p,q) = (0,0) \end{cases}$$

and let $I^r_{p,*} = \sum_q I^r_{p,q}$.

Let $\{D^r\}$ and $\{{}^iD^r\}$, $1 \leqq i \leqq k$, be the spectral sequences of the fibre spaces, (PT_0, π, T_0) and $(PS^{n_i}, \pi_i, S^{n_i})$ respectively.

The spectral sequence $\{{}^iD^r\}$ is easily described by the following diagram.



${}^iD^r_{p,0}$

$$\text{i.e. } {}^i\bar{D}^2_{p,0} \approx \begin{cases} 0 & \text{if } p \neq n_i \\ Z & \text{if } p = n_i \end{cases}$$

$${}^iD^2_{0,*} \approx Z[\alpha_i], \ \dim \alpha_i = n_i - 1$$

$${}^iD^2 \approx \cdots \approx {}^iD^{n_i} \text{ and } {}^i\bar{D}^r = 0, \ r > n_i.$$

Let $F_0({}^iD^r) = \{x : x \in {}^iD^r \text{ and either } x \in {}^iD^r_{0,*} \text{ or } x = 0\}$, and let $F_1({}^iD^r) = {}^iD^r$. This defines a filtration on the differential group ${}^iD^r$.

It follows from [3] that the spectral sequences D^r and ${}^1D^r \otimes \cdots \otimes {}^kD^r$ are isomorphic. Using the above filtrations, define

$$F_i(D^r) = \bigcup_{\Sigma j_i = i} F_{j_1}({}^1D^r) \otimes \cdots \otimes F_{j_k}({}^kD^r)$$

Clearly $F_0(D^r) \subset F_1(D^r) \subset \cdots \subset F_k(D^r) = D^r$. Since D^r may be considered as a tensor product, $d_D{}^r(F_i(D^r)) \subset F_{i-1}(D^r)$ and thus $\{F_1\}$ is a filtration on D^r. We remark that if $y \in F_i(D^r)$ is a cycle and $[y]$ its class, then $[y] \in F_i(D^{r+1})$.

THEOREM 3.3. *There exists a spectral sequence $\{I^r, d_I{}^r\}$, $r \geqq 2$, a map $i^r : I^r \to D^r$ and a family of subgroups of I^r, $\{\mathfrak{F}_i(I^r)\}$ such that*

1. *i^r maps $\mathfrak{F}_j(I^r)$ into $F_j(D^r)$, $i^r \mid \mathfrak{F}_{k-2}(I^r)$ is an isomorphism onto $F_{k-2}(D^r)$ and $i^r \mid \mathfrak{F}_{k-1}(I^r)$ is an epimorphism onto $F_{k-1}(D^r)$.*

2. *$i^r d_I{}^r = d_D{}^r i^r$.*

3. *$\{\mathfrak{F}_i(I^r)\}$ is a filtration on I^r.*

Proof. Let $I^2 = F_{k-1}(D^2)$, i^2 be inclusion, $d_I{}^2 = d_D{}^2$ and $\mathfrak{F}_j(I^2) = F_j(D^2)$ for $j < k$. Assume inductively that I^r, $d_I{}^r$, i^r, and $\mathfrak{F}_i$ have been defined for $r < n$ and satisfy conditions 1-3.

I^n is defined to be the homology of I^{n-1} in the usual sense of spectral sequences. $\mathcal{F}_j(I^n)$ and i^n are induced by $\mathcal{F}_j(I^{n-1})$ and i^{n-1} in the obvious way. Clearly $i^n\colon \mathcal{F}_j(I^n) \to F_j(D^n)$.

Assume that property 1 has been shown to hold for n. Since $I^n = \mathcal{F}_{k-1}(I^n)$, $d_D{}^n i^n(I^n) \subset F_{k-2}(D^n)$. Define $d_I{}^n = (i^n)^{-1} d_D{}^n i^n$. This is well-defined since $i^n \mid \mathcal{F}_{k-2}(I^n)$ is an isomorphism. Properties 2 and 3 follow immediately from this definition. The proof that property 1 is satisfied is straightforward and is left to the reader.

Let X be a simply connected space and let $\{E^r\}$ be the spectral sequence of the fibre space (PX, π, X). We note that $E^2{}_{0,*} \approx H_*(\Omega X)$ is a ring under the Pontryagin product. Bott and Samelson [1] have defined a map $E^r{}_{p,q} \otimes E^2{}_{0,q'} \to E^r{}_{p,q+q'}$, called the action of $E^2{}_{0,q'}$ on E^r. This action is defined inductively as follows. Let $w \otimes \eta$ be a generator of E^2 and let $\gamma \in E^2{}_{0,q'}$. Define $(w \otimes \eta) \star \gamma = w \otimes (\eta\gamma)$ where $\eta\gamma$ is the Pontryagin product of η and γ. This is extended by linearity to E^2. It is easily verified that $d_E{}^2(x \star \gamma) = (d_E{}^2 x) \star \gamma$. Assume inductively that the action has been defined on E^{n-1} and that $d_E{}^{n-1}(x \star \gamma) = (d_E{}^{n-1} x) \star \gamma$. Let w be a cycle in E^{n-1} and $[w]$ its class in E^n. Define $[w] \star \gamma = [w \star \gamma]$ for $\gamma \in E^2{}_{0,*}$. It is shown in [1] that this is well-defined and that $d_E{}^n([w] \star \gamma) = (d_E{}^n[w]) \star \gamma$.

Let $\{E^r\}$ and $\{'E^r\}$ be spectral sequences with actions defined as above. A spectral sequence map $\{h^r\}\colon \{E^r\} \to \{'E^r\}$ is called a homomorphism with respect to $\star$ if $h^r(x \star \gamma) = h^r(x) \star h^2(\gamma)$ for all $x \in E^r$, $\gamma \in E^2{}_{0,*}$.

The action of $D^2{}_{0,q}$ on D^r induces an action of $I^2{}_{0,q}$ on I^r. It is clear that $\{i^r\}$ is a homomorphism relative to these actions.

Calculation of I^∞. If w is a cycle in I^r, we shall not distinguish notationally between w and its homology class in I^{r+1}. Let

$$y = a_1 \otimes \cdots \otimes a_k \in D^2{}_{N,0}.$$

Since $d_D{}^r = 0$ for $r < n_1$ † it follows that y may be considered as an element of $D^{n_1}{}_{N,0}$. A direct calculation shows $d_D{}^{n_1} y \neq 0$ in $D^{n_1}{}_{N-n_1, n_1-1}$. Since $F_{k-1}(D^{n_1}) \approx F_{k-1}(D^2) \approx I^2$, there exists a unique $x \in I^2{}_{N-n_1,n_1-1}$ such that $i^{n_1} x = d_D{}^{n_1} y$. ($x \in I^{n_1}$ since $d_I{}^r = 0$, $r < n_1$).

LEMMA 3.4. *$d_I{}^r(x) = 0$ for all $r \geqq 2$.*

Proof. Since $d_I{}^r(x) = (i^r)^{-1} d_D{}^r i^r(x)$ it suffices to show that $d_D{}^r i^r(x) = 0$. We have shown this above if $r < n_1$. If $r = n_1$, $d_D{}^{n_1} i^{n_1}(x) = d_D{}^{n_1} d_D{}^{n_1} y = 0$. For $r > n_1$ we have $i^r(x) = 0$, since $i^{n_1}(x)$ is a boundary.

COROLLARY. *$x \star \gamma$ is a cycle for all $\gamma \in I^2{}_{0,*}$ and for all $r \geqq 2$.*

---

† Recall that we have assumed $n_1 \leqq n_2 \leqq \cdots \leqq n_k$.

Since x is an element of highest filtration in I^r, it is never a boundary. Therefore $x \star \gamma$ is non-zero in I^∞ for each non-zero $\gamma \in I^2{}_{0,*}$.

THEOREM 3.5. *$\bar{I}^\infty \approx \bar{I}^{n_k+1} \approx G_x \star I^2{}_{0,*} \approx G_x \otimes I^2{}_{0,*}$ where G_x is the free group generated by x.*

Proof. We have shown above that $G_x \star I^2{}_{0,*}$ is contained in $\bar{I}^\infty$. We give an outline of the proof of the other inclusion. If $z \in \bar{I}^\infty$ is non-zero, it must have filtration $k-1$, since $\mathcal{F}_{k-2}(\bar{I}^{n_k+1}) \approx F_{k-2}(\bar{D}^{n_k+1}) = 0$. Therefore, since $\bar{D}^\infty = 0$, $i^r(z)$ is the boundary of an element of filtration k in D^r, for some r. The generators of filtration k in D^2 are of the form $y \otimes \gamma$ for $\gamma \in D^2{}_{0,*}$. Therefore $i^{n_1}(z) = d^{n_1}(y \otimes \gamma) = (d^{n_1}y) \star \gamma = x \star \gamma$. Since i^{n_1} is a monomorphism, $z = x \star \gamma$. Therefore $\bar{I}^\infty \approx G_x \star I^2{}_{0,*}$. The other isomorphisms follows trivially.

Let G_β be the free group generated by β. Let the spectral sequence $\{K^r\}$ be defined as follows.

$$K^2{}_{p,q} = I^2{}_{p,q} \text{ for } p > 0$$
$$K^2{}_{0,q} = I^2{}_{0,q} \oplus G_\beta \otimes I^2{}_{0,q-N+2}$$

and $d_K{}^r = \begin{Bmatrix} d_I{}^r \text{ on the elements of } I^r{}_{p,q} \\ 0 \text{ on } G_\beta \otimes I^2{}_{0,q-N+2} \end{Bmatrix}$ for $r < N - n_1$.

Theorem 3.5 implies that

$$\bar{K}^{N-n_1} \approx G_x \otimes I^2{}_{0,*} \oplus G_\beta \otimes I^2{}_{0,*}.$$

Define $d_K{}^{N-n_1}(x \otimes \gamma) = \beta \otimes \gamma$, for $\gamma \in I^2{}_{0,*}$. It is clear that $\{K^r, d_K{}^r\}$ is a spectral sequence and $\bar{K}^\infty = 0$.

The spectral sequence $\{E^r\}$ desired in Theorem 3.2 will be defined as a direct sum of copies of K^∞. Before defining $\{E^r\}$ we must define the set over which the direct sum is to be taken.

Let $R = Z[\alpha_1, \cdots, \alpha_k] * Z[\beta]$. R is graded by assigning each element of R a dimension as follows.

$$\dim e = 0 \qquad (e \text{ is the identity in } R).$$
$$\dim \alpha_i = n_i - 1 \qquad i = 1, \cdots, k$$
$$\dim \beta = N - 2 \qquad N = \sum_{i=1}^{k} n_i$$
$$\dim (r_1 r_2) = \dim r_1 + \dim r_2$$

Denote the set of elements of dimension q by R_q. (Our goal is to show that $H_*(\Omega T_1) \approx R$.)

R_q is a Z-module freely generated by a set of monomials in the α_i's and

β for $q > 0$, and by e, when $q = 0$. Let V_q be a set of these generators and set $V = \bigcup_q V_q$.

Let T be the right ideal of R generated by β and let P be the positive dimensional elements of $Z[\alpha_1, \cdots, \alpha_k]$. We state a lemma which we will need later.

LEMMA 3.6.

a) $R \approx R_0 + P + T + P \otimes T$

b) *βV is a set of generators for T.*

For each $v \in V$ let $G_{\beta v}$ be the free group generated by βv. If $v \in V_t$, define

$$^vE^r{}_{p,q} = \tilde{K}^r{}_{p,q-t-N+2} \otimes G_{\beta v}.$$

We write $^vE^r \approx \tilde{K}^r \otimes G_{\beta v}$ keeping in mind the shift in dimension. It is then obvious that $\{^vE^r, d_K{}^r \otimes 1\}$ is a spectral sequence. Since $\tilde{K}^\infty = 0$, $^vE^\infty = 0$.

Definition 3.7. $E^r = K^r \otimes G_e \oplus \sum_{v \in V} {}^vE^r$ and $d_E{}^r = d_K{}^r \otimes 1$.

Since a direct sum of spectral sequences is a spectral sequence, we have the following lemma.

LEMMA 3.8. *$\{E^r, d_E{}^r\}$ is a spectral sequence and $\bar{E}^\infty = 0$.*

It will be shown below that $\{E^r\}$ is the spectral sequence desired in Theorem 3.2. The definition of $\{E^r\}$ was chosen to make the proof of property 4 trivial. We must show that properties 1-3 hold.

LEMMA 3.9. *$E^2{}_{0,*}$ may be given a ring structure such that the rings $E^2{}_{0,*}$ and R are isomorphic.*

Proof. We first show that $E^2{}_{0,*}$ and R are isomorphic as modules. The proof uses Lemma 3.6 and the fact that $I^2{}_{0,*} \approx P + R_0$.

$$\begin{aligned}
E^2{}_{0,*} &= K^2{}_{0,*} \otimes G_e + \sum_{v \in V} \tilde{K}^2{}_{0,*} \otimes G_{\beta v} \\
&\approx I^2{}_{0,*} + G_\beta \otimes I^2{}_{0,*} + \sum_{v \in V} (\tilde{I}^2{}_{0,*} \otimes G_{\beta v} + G_\beta \otimes I^2{}_{0,*} \otimes G_{\beta v}) \\
&\approx R_0 + P + G_\beta \otimes R_0 + G_\beta \otimes P + P \otimes T + G_\beta \otimes T + G_\beta \otimes P \otimes T \\
&\approx R_0 + P + P \otimes T + G_\beta \otimes (R_0 + P + T + P \otimes T) \\
&\approx R_0 + P + P \otimes T + T \\
&\approx R.
\end{aligned}$$

Let f be the map giving the above isomorphism. It is clear that $f(x \otimes y) = xy$.

Define a ring structure on $E^2{}_{0,*}$ by setting

$$(u_1 \otimes t_1)(u_2 \otimes t_2) = \begin{cases} u_1 \otimes t_1 f(u_2 \otimes t_2) & \text{if } \begin{cases} t_1 \in T \text{ or if } t_1 = e \text{ and} \\ u_2 \in G_\beta \otimes I^2{}_{0,q} \end{cases} \\ (u_1 \star u_2) \otimes t_2 & \text{if } t_1 = e \text{ and } u_2 \in \bar{I}^2{}_{0,q} \end{cases}$$

This is well defined since either $t_1 f(u_2 \otimes t_2) \in T$, $t_2 \in T$, or $t_2 = e$. If $t_1 = t_2 = e$ and $u_1, u_2 \in \bar{I}^2{}_{0,q}$, this is the usual multiplication in $I^2{}_{0,*}$.

In either case a direct calculation shows that f is a ring homomorphism. Since f is a module isomorphism, it follows that f is a ring isomorphism.

LEMMA 3.10. $E^2{}_{p,0} \approx H_p(T_1)$.

Proof. Since ${}^vE^2{}_{p,0} = 0$ for all v, $E^2{}_{p,0} \approx K^2{}_{p,0} \approx I^2{}_{p,0} \approx H_p(T_1)$.

LEMMA 3.11. $E^2{}_{p,q} \approx E^2{}_{p,0} \otimes E^2{}_{0,q}$.

Proof. We may assume that p and q are positive.

$$\begin{aligned} E^2{}_{p,q} &\approx K^2{}_{p,q} \oplus \sum_{v \in V} I^2{}_{p,q-\dim(\beta v)} \otimes G_{\beta v} \\ &\approx I^2{}_{p,q} \oplus \sum_{v \in V} I^2{}_{p,q-\dim(\beta v)} \otimes G_{\beta v} \\ &\approx I^2{}_{p,0} \otimes (I^2{}_{0,q} \oplus \sum_{v \in V} I^2{}_{0,q-\dim(\beta v)} \otimes G_{\beta v}) \\ &\approx E^2{}_{p,0} \otimes E^2{}_{0,q}. \end{aligned}$$

Since $\{E^r\}$ satisfies properties 1-4, we have completed the proof of Theorem 3.2.

Construction of the homomorphism $\{h^r\}: \{E^r\} \to \{A^r\}$.

To conclude the proof of Theorem 2.1 we must prove:

THEOREM 3.12. *If* $\{A^r\}$ *is the spectral sequence of the fibre space* (PT_1, π, T_1), *there exists a map* $\{h^r\}: \{E^r\} \to \{A^r\}$ *such that*

1. h^r *is a homomorphism of spectral sequences*
2. $h^2{}_{p,0}$ *is an isomorphism for all* p.
3. $h^2{}_{0,q}$ *is a ring homomorphism.*
4. h^∞ *is an isomorphism.*

Our procedure is to define a map of spectral sequences $\{\lambda^r\}: \{K^r\} \to \{A^r\}$ and then define h^r using the fact that E^r is a direct sum. We first study $\{A^r\}$.

LEMMA 3.13. $H_*(\Omega T_1)$ *has a subring isomorphic to* $H_*(\Omega T_0)$.

Proof. Let $j: T_1 \to T_0$ and $l: T_{k-1} \to T_1$ be inclusions.

$$(\Omega(jl))_* H_*(\Omega T_{k-1}) \to H_*(\Omega T_0)$$

is an epimorphism. For $k > 2$, the commutators in $H_*(\Omega T_{k-1})$ are in the kernel of $(\Omega l)_*$ since the corresponding Whitehead products are in the kernel of $l_*: \pi_*(T_{k-1}) \to \pi_*(T_1)$.

Remark. Lemma 3.13 is true with $H_*(\Omega T_i)$, $i < k-1$, replacing $H_*(\Omega T_1)$.

We denote this subring of $H_*(\Omega T_1)$ by $Z[\zeta_1, \cdots, \zeta_k]$, where $\dim \zeta_i = n_i - 1$. The generators of $A^2_{*,0}$ which transgress to the ζ_i's are denoted by $z_1, \cdots, z_k$, $\dim z_i = n_i$. The remaining generators of $\tilde{A}^2_{*,0} \approx H_*(T_1)$ are tensor products of the z_i's.

Let $\{j^r\} : \{A^r\} \to \{D^r\}$ be the map induced by inclusion. There is an action of $A^2_{0,*}$ on A^r defined as usual, and j^r is a homomorphism relative to this action. Clearly $j^2(\zeta_i) = \alpha_i$ and $j^2(z_i) = a_i$. We now show that there is an element $\eta \in A^2_{0,N-2}$ corresponding to the element $\beta \in E^2_{0,N-2}$.

Since $j^2_{p,0}$ is an isomorphism for $p < N$, $j^2_{0,q}$ is an isomorphism for $q < N-2$ by Theorem 3.1. Therefore $j^2_{p,q}$ is an isomorphism for $p < N$ and $q < N-2$. In particular there is a unique element, $u \in A^2_{N-n_1,n_1-1}$ such that $j^2(u) = i^2(x)$.

LEMMA 3.14.

i) *u is a non-zero cycle in A^r for $r < N - n_1$*

ii) $d_A^{N-n_1} u \neq 0$.

Proof. i) $j^r d_A^r(u) = d_D^r j^r(u)$. Since

$$d_D^r = 0 \text{ if } r < n_1$$
$$d_D^{n_1} j^{n_1} u = d_D^{n_1} d_D^{n_1} y = 0, \text{ and}$$
$$j^r(u) = i^r(x) = 0 \text{ if } r > n_1,$$

$j^r d_A^r(u) = 0$ for all r.

By Theorem 3.1, $j^r_{N-n_1-r,n_1+r-2}$ is a monomorphism for $r < N - n_1$. Therefore $d_A^r u = 0$ if $r < N - n_1$. We may define a filtration on A^r in a manner similar to the filtration defined on D^r. Since u is an element of highest filtration it can never be a boundary. Hence $u \neq 0$ in A^r, $r \leqq N - n_1$.

ii) $d_A^{N-n_1}(u)$ is non-zero. For if it were zero, u would be a non-zero element of $\tilde{A}^\infty$ which is impossible.

Let $\eta \in A^2_{0,N-2}$ be chosen so that $d_A^{N-n_1} u = \eta$. Since the class of each element in $Z[\zeta_1, \cdots, \zeta_k]$ is zero in A^{n_k+1}, η may be chosen to be independent of $Z[\zeta_1, \cdots, \zeta_k]$.

For each $\gamma \in A^2_{0,q}$ we then have

$$d_A^r(u \star \gamma) = \begin{cases} 0 \text{ if } r < N - n_1 \\ \eta \star \gamma \text{ if } r = N - n_1. \end{cases}$$

Define $\{\lambda^r\} : \{K^r\} \to \{A^r\}$ inductively as follows. $\lambda^2 : K^2 \to A^2$ is given by setting $\lambda^2(a_i) = z_i$, $\lambda^2(\alpha_i) = \zeta_i$, $\lambda^2(\beta) = \eta$ and extending to K^2 using

the ring and tensor product structures of K^2. Assume λ^r has been defined for $r < n$. We would like to define λ^n to be the map induced in homology by λ^{n-1}. For such a map to be well-defined $d_A{}^{n-1}\lambda^{n-1}$ must be equal to $\lambda^{n-1}d_K{}^{n-1}$. The proof that this is indeed the case is, unfortunately, complicated. We prove this by induction on n.

To show that $\lambda^2 d_K{}^2 = d_A{}^2\lambda^2$ it suffices to show that for each generator $w \in K^2{}_{p,0}$, $\lambda^2 d_K{}^2 w = d_A{}^2\lambda^2 w$.

For each m, $1 \leqq m \leqq k$, let $T_0{}^m$ equal $T_0(S^{n_1}, \cdots, S^{n_{m-1}}, S^{n_{m+1}}, \cdots, S^{n_k})$ and let $\{J_m{}^r\}$ be the spectral sequence associated with the fibre space $(PT_0{}^m, \pi_m, T_0{}^m)$. The canonical injection $\phi_m\colon T_0{}^m \to T_1$, induces a homomorphism of spectral sequence $\{\phi_m{}^r\}\colon \{J_m{}^r\} \to \{A^r\}$.

The image of $J_m{}^2$ in D^2 under injection is contained in $F_{k-1}(D^2) = I^2 \subset K^2$. This induces a homomorphism of spectral sequences $\{\theta_m{}^r\}\colon \{J_m{}^r\} \to \{K^r\}$ such that $\lambda^2\theta_m{}^2 = \phi_m{}^2$. We would like to write $\lambda^r\theta_m{}^r = \phi_m{}^r$, however to do this we must first show that λ^r is well-defined for $r > 2$.

Each generator $w \in K^2{}_{*,0}$ is in the image of $\theta_m{}^2$ for some m. Let $\theta_m{}^2 v = w$. Then

$$\begin{aligned} d_A{}^2\lambda^2(w) &= d_A{}^2\lambda^2\theta_m{}^2(v) = d_A{}^2\phi_m{}^2(v) \\ &= \phi_m{}^2 d_{J_m}{}^2(v) = \lambda^2\theta_m{}^2 d_{J_m}{}^2(v) \\ &= \lambda^2 d_K{}^2\theta_m{}^2(v) = \lambda^2 d_K{}^2(w). \end{aligned}$$

Assume inductively that for $r < n$

1. λ^r is defined.
2. $d_A{}^r\lambda^r = d_K{}^r\lambda^r$.
3. $\lambda^r\theta_m{}^r = \phi_m{}^r$.

Define λ^n to be the map induced in homology by λ^{n-1}. This is easily seen to be well-defined and $\lambda^n\theta_m{}^n = \phi_m{}^n$.

The generator of K^n are of three types: $w \star \gamma$, $x \star \gamma$, and $\beta \otimes \gamma$ where $w \in I^n{}_{p,0}$ and $\gamma \in I^2{}_{0,q}$. For $n \leqq n_1$ types 1 and 2 intersect. A direct calculation shows that $d_A{}^n\lambda^n = \lambda^n d_K{}^n$ on generators of the second and third type. A calculation similar to the one used when $n = 2$ proves the result on generators of type 1.

Thus we have shown that $\{\lambda^r\}\colon \{K^r\} \to \{A^r\}$ is a homomorphism of spectral sequences.

Define $\mu\colon R \to A^2$ by setting $\mu(\alpha_i) = \zeta_i$, $\mu(\beta) = \eta$, $\mu(e) = e'$ and extending to a ring homomorphism (e is the generator of $A^2{}_{0,0}$). $\{h^r\}\colon \{E^r\} \to \{A^r\}$ is defined by

$$h^r(w \otimes r) = \lambda^r(w) \star \mu(r)$$

for $w \in K^r$ and $r \in V$ or $r = e$. It follows that h^r is a homomorphism of spectral sequences, since

$$\begin{aligned} d_A{}^r h^r(w \otimes r) &= d_A{}^r(\lambda^r(w) \star \mu(r)) \\ &= (d_A{}^r \lambda^r(w)) \star \mu(r) \\ &= (\lambda^r d_K{}^r(w)) \star \mu(r) \\ &= h^r(d_K{}^r(w) \otimes r) \\ &= h^r d_B{}^r(w \otimes r). \end{aligned}$$

Since $E^2{}_{p,0} = K^2{}_{p,0} = I^2{}_{p,0}$, $h^2 \mid E^2{}_{p,0} = \lambda^2 \mid K^2{}_{p,0}$ which is an isomorphism. Thus property 2 is satisfied. Property 4 follows since $\bar{E}^\infty \cong \bar{A}^\infty = 0$. Since $\lambda^2(\gamma) = \mu(\gamma)$ for $\gamma \in K^2{}_{0,q} \subset R$, property 3 is easily checked. This completes the proof of Theorem 3.12.

Proof of Theorem 2.1. Theorem 3.2 and 3.12 imply that the hypothesis of Theorem 3.1 are satisfied for $n = \infty$. Thus $A^2{}_{0,*} \approx E^2{}_{0,*}$ as a ring. Lemma 3.9 then implies

$$H_*(\Omega T_1) \approx A^2{}_{0,*} \approx R \approx H_*(\Omega T_0) * H_*(\Omega S^{N-2}).$$

This completes the proof of Theorem 2.1.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] R. Bott and H. Samelson, "On the Pontryagin product in spaces of paths," *Commentarii Mathematici Helvetici*, vol. 27 (1953), pp. 320-337.

[2] E. Dyer and L. K. Lashof, "Homology of iterated loop spaces," *American Journal of Mathematics*, vol. 84 (1962), pp. 35-88.

[3] V. K. A. M. Gugenheim and J. C. Moore, "Acyclic models and fibre spaces," *Transactions of the American Mathematical Society*, vol. 85 (1957), pp. 265-306.

[4] P. J. Hilton, "On the homotopy groups of the union of spheres," *Journal of the London Mathematical Society*, vol. 30 (1955), pp. 154-172.

[5] J. Milnor, *The construction FK*, Princeton University (1956), mimeographed.

[6] G. J. Porter, "Higher order Whitehead products," *Topology* (to appear).

[7] H. Samelson, "A connection between the Whitehead and the Pontryagin product," *American Journal of Mathematics*, vol. 75 (1953), pp. 744-752.

[8] J. P. Serre, "Homologie singuliere des espaces fibres." *Annals of Mathematics*, (2), vol. 54 (1951), pp. 425-505.

[9] E. C. Zeeman, "A proof of the comparison theorem for spectral sequences," *Proceedings of the Cambridge Philosophical Society*, vol. 53 (1957), pp. 57-62.

NON-LOCAL BOUNDARY VALUE PROBLEMS FOR ELLIPTIC OPERATORS.

By Richard Beals.

Introduction. A number of investigations have been made in recent years of boundary value problems for elliptic partial differential operators. In most cases the boundary conditions considered have been *local,* i. e. determined by differential operators; see, for example, [1], [7], [8], [9], [22], [24], [25]. Recently Bade and Freeman [3] have studied a class of boundary value problems for the Laplace operator in which the boundary operators are more general. This work has been extended by Freeman [17] to general strongly elliptic formally self-adjoint second order operators. In the present paper we study more general classes of non-local boundary value problems for elliptic operators of higher order. The approach is similar in outline to that in [3], but rather different analytical methods are used. These results have been announced in [6].

This paper is part of the author's doctoral dissertation, prepared under the direction of Professor Felix Browder and presented for the degree of Doctor of Philosophy at Yale University. The author is grateful to Professor Browder for suggesting the problem, and for his advice and encouragement throughout. The research was done while the author held a National Science Foundation Cooperative Fellowship.

Summary and outline. We begin by indicating the line of approach and some of the main results.

For second order formally self adjoint operators the basic idea is due to Calkin [13], [14], and is carried out for the Laplace operator in [3]. Let G be an open set in n-space having smooth boundary Γ, and let A be the Laplace operator $(\partial/\partial x_1)^2 + \cdots + (\partial/\partial x_n)^2$. One of the usual Green's formulas can be looked upon as expressing the fact that a certain operator S acting in $L^2(G) \oplus L^2(\Gamma)$ is symmetric, and if the domain of S is chosen properly it is self-adjoint. There is a duality between properties of perturbaitons of S and properties of realizations of A as an operator in $L^2(G)$ corresponding to appropriate boundary conditions. These properties and the

Received June 2, 1964.
Revised September 21, 1964.

duality are established in [3] by means of potential theory. In a sense the operator S is the natural object to consider, incorporating both A and the boundary conditions. The central result of this paper is the construction of the analogous operator S for a general elliptic operator. This construction is made possible by certain a priori estimates which replace the potential theory of [3].

All operators considered are linear. We denote the domain and range of an operator T by $D(T)$ and $R(T)$ respectively. Let A be an elliptic of order $2p$ defined on the closure of an open set G having smooth boundary Γ. Let $B=(B_0,B_1,\cdots,B_{2p-1})$ be a system of differential operators defined near Γ, with each B_k an operator of order k for which the surface Γ is non-characteristic. Then by [2, Theorem VI] there are operators A', B' also satisfying these conditions and which are *conjugate* to A, B in the sense that for any smooth functions u and v defined on $G\cup\Gamma$ and having compact support, we have the Green's formula

$$(Au,v)-(u,A'v)=\sum(-1)^{k+1}(B_{2p-1-k}u,B_k'v). \tag{1}$$

Here the parentheses on left and right denote inner products in $L^2(G)$ and $L^2(\Gamma)$ respectively.

The associated operators S and S' exist under the following conditions. We assume that Λ_0 and Λ_1 are complementary subsets of $\{0,1,\cdots,2p-1\}$ having p elements in decreasing and in increasing order respectively; set

$$\Lambda_0'=(2p-1-j\mid j\in\Lambda_1) \text{ and } \Lambda_1'=(2p-1-j\mid j\in\Lambda_0).$$

We assume: (i) G has sufficiently smooth boundary; (ii) the coefficients of the operators are sufficiently smooth; (iii) B_k and B_k' are of order k, with Γ uniformly non-characteristic for each; (iv) A and A' are uniformly and regularly elliptic; (v) the systems $(A,(B_k\mid k\in\Lambda_1))$ and $(A',(B_k'\mid k\in\Lambda_1'))$ are "regular," or satisfy the "root covering condition," uniformly; (vi) equation (1) holds. The precise form of these conditions is given in Chapter 4, (S1)-(S6).

For s a non-negative integer, the Sobolev space $W^{s,2}(G)$ is the space of distributions on G which have all derivations of order $\leq s$ in $L^2(G)$. The spaces $W^{s,2}(\Gamma)$ and also the spaces $W^{s,2}(G)$ for s not an integer are defined in Chapter 1, as are "L^2-boundary values." An operator $T\colon X\to Y$, where X and Y are Banach spaces, is said to be a *Fredholm operator* if T is closed, $N(T)$ has finite dimension, and $R(T)$ is closed and has finite co-dimension in Y.

With these preliminaries we can now state

THEOREM *A*. *Suppose G, A, B, A', and B' satisfy conditions* (S1)-(S6) *of Chapter 4 below, with respect to the index sets Λ_0, Λ_1, Λ_0', Λ_1'. Let S_0 be the operator mapping $L^2(G) \oplus E_1$ into $L^2(G) \oplus E_2$, where*

$$E_1 = \sum_{j \in \Lambda_0} \oplus W^{-j,2}(\Gamma), \qquad E_2 = \sum_{j \in \Lambda_1} \oplus W^{j,2}(\Gamma),$$

having as domain all $[u, ((-1)^j B_j u \mid j \in \Lambda_0)]$ for u in $C_0^\infty(G \cup \Gamma)$ and with $S_0[u, ((-1)^j B_j u)] = [Au, (B_j u \mid j \in \Lambda_1)]$. Let S_0' be defined similarly. Let Ω be the set of all u in $L^2(G)$ such that Au is in $L^2(G)$ and such that the distribution derivatives of u of order $< 2p$ have L^2-boundary values. Let Ω' be defined similarly. Then:

(a) *S_0 and S_0' are closable, and their closures S and S' are adjoints.*

(b) *$D(S) = \{[u, (-1)^j B_j u] \mid u \in \Omega\}$ and similarly for $D(S')$.*

(c) *Ω and Ω' are contained in $W^{2p-\frac{1}{2},2}(G)$.*

(d) *If G is bounded, then S and S' are Fredholm operators.*

(Note that it follows immediately from (1) and the definitions that $S_0 \subseteq (S_0')^*$ and $S_0' \subseteq S_0^*$.)

Next we wish to consider boundary value problems. For convenience we define boundary operators by

$$\gamma_0 = ((-1)^j B_j \mid j \in \Lambda_0), \qquad \gamma_1 = (B_k \mid k \in \Lambda_1). \tag{2}$$

Operators γ_0' and γ_1' are defined similarly, and spaces E_1' and E_2' are defined as above. Then S is given by $S[u, \gamma_0 u] = [Au, \gamma_1 u]$. Associated with S is the operator $A(0)$, the restriction of A to those u in Ω such that $\gamma_1 u = 0$.

Now suppose $C: E_1 \to E_2$. Define the *induced operator* $C_1: (L^2(G) \oplus E_1) \to (L^2(G) \oplus E_2)$ by setting $C_1[g, f] = [0, Cf]$. Associated with the perturbed operator $S - C_1$ is $A(C)$, the restriction of A to those u in Ω such that $(\gamma_1 - C\gamma_0)u = 0$. This corresponds to the boundary value problem

$$Au = f \text{ in } G, \qquad \gamma_1 u = C\gamma_0 u \text{ on } \Gamma. \tag{*}$$

Exploiting the duality between $S - C_1$ and $A(C)$, we obtain the following.

THEOREM *B*. *Let G, A, B, A', and B' satisfy conditions* (S1)-(S6) *of Chapter 4. Let γ_0 and γ_1 be defined by* (2), *and let E_1, E_2, and Ω be defined as in Theorem A. Let $C: E_1 \to E_2$ be such that $D(C) \supseteq \gamma_0(\Omega)$ and such that for some positive constants $\epsilon < 1$ and K the inequality*

$$\| C\gamma_0 u \|_{E_1} \leqq \epsilon(\| Au \|_{L^2(G)} + \| \gamma_1 u \|_{E_2}) + K \| u \|_{L^2(G)}$$

holds for all u in Ω. Let $A(C)$ be the restriction of A to those u in Ω such that $\gamma_1 u = C\gamma_0 u$. Then:

(a) *$A(C)$ is a closed operator in $L^2(G)$.*

(b) *If in addition $R(C) \subseteq \sum W^{k+\frac{1}{2},2}(\Gamma)$ $(k \in \Lambda_1)$, then the domain of $A(C)$ is contained in $W^{2p,2}(G)$.*

Let X and Y be Banach spaces, and $T: X \rightarrow Y$. An operator $C: X \rightarrow Y$ is said to be *T-compact* if $D(C) \supseteq D(T)$ and C is compact as a mapping from $D(T)$ to Y with respect to the graph topology on $D(T)$.

THEOREM C. *Let G, A, B, A', and B' satisfy conditions* (S1)-(S6) *of Chapter* 4. *Let operators $S, \gamma_0, \cdots$, and spaces $E_1, \cdots$, be as in Theorems A and B. Suppose $C: E_1 \rightarrow E_2$ and $C': E_1' \rightarrow E_2'$ are such that the induced operators C_1 and C_1' are closable and S-compact and S'-compact respectively. Let $A(C)$ be the restriction of A to those u in Ω such that $\gamma_1 u = C\gamma_0 u$, and let $A'(C')$ be defined similarly. Then:*

(a) *$A(C)$ and $A'(C')$ are closed in $L^2(G)$.*

(b) *If $(S - C_1)^* = S' - C_1'$, then $A(C)$ and $A'(C')$ are adjoints.*

(c) *If G is bounded, then $A(C)$ and $A'(C')$ are Fredholm operators with the same indices as S and S' respectively. Moreover if $C' \subseteq C^*$, then the condition in* (b) *is necessarily satisfied.*

In Chapter 1 below we introduce the various function spaces and norms to be used throughout. In Chapter 2 we set forth the analytical machinery replacing the potential theory of [3]: (a) an a priori estimate on the $W^{2p-\frac{1}{2},2}(G)$ norm of u in terms of the L^2-norm of Au and the E_2-norm of $\gamma_1 u$; (b) a similar estimate on uniform convergence in L^2 of the restrictions of derivatives of u to surfaces "parallel" to Γ; (c) a result on solvability of a non-homogeneous boundary value problem. In Chapter 3 we collect for reference results of Aronszajn and Milgram [2] on existence of conjugate systems, of Browder on a sufficient condition for regularity of (A, γ_1), and of Schechter [25] on regularity of adjoint systems.

Chapter 4 is devoted to deriving Theorem A (a) and (b), using the results of Chapter 2 together with known facts about $A(0)$. In Chapter 5 we develop abstractly the connections between properties of S and properties of the corresponding realizations of A. These are applied to non-local boundary value problems, obtaining Theorem B, Theorem C (a) and (b), and other results. In Chapter 6 we consider the sharper results, such as Theorem A(c) and Theorem C(c), obtainable when G is bounded, using the Sobolev imbedding theorem, a result of Browder [9, Lemma 1], the theory of Fredholm operators, and a remark of the author [5].

Finally, Chapter 7 is devoted to second order operators, chiefly to variations on the Neumann problem for strongly elliptic operators. Here we extend the results of [3] on the location of the spectrum of the realizations $A(C)$, in particular for any bounded operator C in $L^2(\Gamma)$.

Chapter 1. Function Spaces on Regular Domains.

We denote by E^n n-dimensional Euclidean space, with Lebesque measure. If $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ is an n-tuple of non-negative integers and $x = (x_1, x_2, \cdots, x_n)$ is in E^n, then $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$. Similarly the differential operator D^α is $D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n}$, where $D_j = (i)^{-1}\partial/\partial x_j$. The *order* of D^α is $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$. If u is a complex-valued measurable function defined on E^n, the *support* of u, denoted $\operatorname{supp}(u)$, is the intersection of all closed sets F such that $u = 0$ almost everywhere on $E^n - F$.

We denote by $C^p(E^n)$ the space of all complex-valued functions on E^n having continuous derivatives of order $\leqq p$, and by $C^\infty(E^n)$ the intersection of all the $C^p(E^n)$, $p \geqq 0$. The subspaces consisting of functions having compact support will be denoted by $C_0^p(E^n)$ and $C_0^\infty(E^n)$ respectively.

If X is a linear topological space, X^* will denote the space of continuous conjugate linear functionals on X. The pairing between elements of X^* and of X will be denoted by $\langle x^*, x\rangle$. In particular $\mathfrak{D} = \mathfrak{D}(E^n) = C_0^\infty(E^n)$ with the usual topology [26], and $\mathfrak{D}^*$ is the conjugate space. If v is a complex-valued function on E^n, let $\bar{v}$ denote the complex conjugate. Let Ju be defined for $u \in \mathfrak{D}^*$ by $\langle Ju, \phi\rangle = \langle u, \bar{\phi}\rangle$, $\phi \in \mathfrak{D}$. Then J is an isomorphism of $\mathfrak{D}^*$ onto $\mathfrak{D}'$, the space of distributions.

Similarly, let $\mathcal{S} = \mathcal{S}(E^n) = \{u \mid u \in C^\infty(E^n), x^\alpha D^\beta u \in L^\infty(E^n), \text{ all } \alpha, \beta\}$, with the usual topology [26]. Then the mapping J defined as above is an isomorphism of $\mathcal{S}^*$ with $\mathcal{S}'$, the space of tempered distributions.

The inner product in $L^2(E^n)$ will be denoted by (u, v):

$$(1) \qquad (u, v) = \int_{E^n} u(x)\bar{v}(x)\,dx.$$

There is a natural injection $j: L^2(E^n) \to \mathcal{S}^*$, given by

$$(2) \qquad \langle j(u), \phi\rangle = (u, \phi), \quad u \in L^2(E^n), \quad \phi \in \mathcal{S}.$$

We shall use this injection to consider $L^2(E^n)$ as a subspace of $\mathcal{S}^*$.

Let $R^n = \{\xi\}$ denote the dual space of E^n. We denote the pairing between $\xi \in R^n$ and $x \in E^n$ simply by $\xi x = x\xi$. The *Fourier transform* of a function $u \in \mathcal{S}(E^n)$ is given by

$$(3) \qquad \mathfrak{F}u(\xi) = (2\pi)^{-n/2} \int_{E^n} e^{-ix\xi} u(x)\,dx.$$

Then $\mathcal{F}$ maps $\mathcal{S}(E^n)$ isomorphically onto $\mathcal{S}(R^n)$. The inverse is $\mathcal{F}^*$, where

$$\mathcal{F}^*v(x) = (2\pi)^{-n/2}\int_{R^n} e^{ix\xi}v(\xi)\,d\xi. \tag{4}$$

By the Plancherel theorem $\mathcal{F}$ can be extended to an isometry of $L^2(E^n)$ onto $L^2(R^n)$, and this in turn can be extended to an isomorphism, also denoted by $\mathcal{F}$, of $\mathcal{S}^*(E^n)$ onto $\mathcal{S}^*(R^n)$ by the formula

$$\langle \mathcal{F}u, v\rangle = \langle u, \mathcal{F}^*v\rangle, \quad u \in \mathcal{S}^*(E^n), \quad v \in \mathcal{S}(R^n). \tag{5}$$

For any real s, the space $W^{s,2} = W^{s,2}(E^n)$ is

$$\{u \mid u \in \mathcal{S}^*, \mathcal{F}u(\xi)(1+|\xi|^2)^{s/2} \in L^2(R^n)\}.$$

This is a Hilbert space relative to the inner product

$$(u,v)^*_s = \int_{R^n}(1+|\xi|^2)^s\mathcal{F}u(\xi)(\mathcal{F}v(\xi))^- d\xi. \tag{6}$$

The associated norm is $\| u \|^*_s$. For each s, $\mathcal{D}$ is dense in $W^{s,2}$ (see [22], Chapter 1).

From the Plancherel theorem it follows that for p a non-negative integer, $W^{p,2}$ is the space of all distributions u having all distribution derivatives $D^\alpha u$ for $|\alpha| \leqq p$ in $L^2(E^n)$. We recall that by definition, $\langle D^\alpha u, \phi\rangle = \langle u, D^\alpha \phi\rangle$, $\phi \in \mathcal{D}$. The norm $\| u \|^*_p$ is equivalent to $\| u \|_p$, where

$$\| u \|_p = \Big(\sum_{|\alpha|\leqq p} \| D^\alpha u \|^2\Big)^{\frac{1}{2}}. \tag{7}$$

Here $\| v \|$ denotes the L^2-norm.

Suppose $0 < s < 1$, and $u \in \mathcal{S}$. Define ${}_0\| u \|_s$ by

$$({}_0\| u \|_s)^2 = \int_{E^n}\int_{E^n} |u(x) - u(y)|^2 \, |x-y|^{-n-2s}\,dxdy. \tag{8}$$

Taking Fourier transforms after a change of variables, we see that

$$({}_0\| u \|_s)^2 = k_s\int_{R^n} |\bar{u}(\xi)|^2 \, |\xi|^{2s}\,ds. \tag{9}$$

(See [22, Chapter 1].) Since $\mathcal{S}$ is dense in $W^{s,2}$, the equality (9) carries over to all $u \in W^{s,2}$. Therefore if p is a non-negative integer and $0 < s < 1$, we have the norm $\| u \|^*_{p+s}$ equivalent to $\| u \|_{p+s}$, with

$$(\| u \|_{p+s})^2 = \sum_{|\alpha|\leqq p} \| D^\alpha u \|^2 + \sum_{|\alpha|=p} ({}_0\| D^\alpha u \|_s)^2. \tag{10}$$

We can now proceed to define analogous function spaces on subsets of E^n.

If F is a measurable subset of E^n and u a measurable function defined on F, let u_1 be the extension to E^n obtained by setting $u_1(x)=0$, $x\notin F$. The *support* of u is then defined to be $\operatorname{supp}(u_1)\cap F$. Now let G be an open set in E^n with boundary Γ, and let Γ_0 be a (possibly empty) measurable subset of Γ. We define $C^p(G\cup\Gamma_0)$ to be the space of complex-valued functions on G whose derivatives of order $\leqq p$ all have continuous extensions to $G\cup\Gamma_0$, and $C^\infty(G\cup\Gamma_0)$ to be the intersection for all $p\geqq 0$. The subspaces of functions with compact support are denoted by $C_0{}^p(G\cup\Gamma_0)$ and $C_0{}^\infty(G\cup\Gamma_0)$. Let $\mathfrak{D}(G)=C_0{}^\infty(G)$ with the usual topology [26], and let $\mathfrak{D}^*(G)=(\mathfrak{D}(G))^*$.

We denote the inner product in $L^2(G)$ by $(u,v)_G$, and the corresponding norm by $\|u\|_G$. If $0<s<1$ and u is a measurable function on G, we define ${}_0\|u\|_{s,G}$ by

$$({}_0\|u\|_{s,G})^2=\int_G\int_G|u(x)-u(y)|^2\,|x-y|^{-n-2s}\,dxdy. \tag{11}$$

For p a non-negative integer and $0<s<1$, let

$$W^{p,2}(G)=\{u\mid u\in\mathfrak{D}^*(G),D^\alpha u\in L^2(G),\ \text{all}\ 0\leqq|\alpha|\leqq p\}.$$
$$W^{p+s,2}(G)=\{u\mid u\in W^{p,2}(G),{}_0\|D^\alpha u\|_s<\infty,\ \text{all}\ |\alpha|=p\}.$$

These spaces are given the norms defined by

$$(\|u\|_{p,G})^2=\sum_{|\alpha|\leqq p}(\|D^\alpha u\|_G)^2. \tag{12}$$

$$(\|u\|_{p+s,G})^2=(\|u\|_{s,G})^2+\sum_{|\alpha|=p}({}_0\|D^\alpha u\|_{s,G})^2. \tag{13}$$

These are Hilbert spaces relative to the appropriate inner products, whose forms are evident from (12) and (13).

By the remarks above these definitions are consistent with our usage for $G=E^n$. For $s<0$ we put $W^{s,2}(G)=(W^{-s,2}(G))^*$; this is also seen to be consistent with the previous.

Let G be an open set in E^n. A function $\phi\colon G\to E^m$ is said to be *of class p* if each coordinate $\phi_1,\cdots,\phi_m$ is in $C^p(G)$. It is said to be *uniformly of class p* if also the derivatives of order $\leqq p$ are all bounded. The open set G in E^n is said to be *regular of class p* if for each point x of the boundary Γ there is a neighborhood N of x and a homeomorphism ϕ of N onto the open unit sphere B^n of E^n such that ϕ and ϕ^{-1} are of class p, while $\phi(N\cap G)=B_+{}^n=\{x\mid x\in B^n,x_1>0\}$. If G is regular of class $p\geqq 2$, then for any $s\geqq 0$, $C_0{}^\infty(G\cup\Gamma)$ is dense in $W^{s,2}(G)$; see [9], [22].

The open set G in E^n with boundary Γ is said to be *uniformly regular of class p* if there is a family of distinct open sets $\{N_k\}$, a family of mappings $\{\phi_k\}$, and integer r such that:

(1) Let $M_k = \phi_k^{-1}(\{x \mid |x| < \frac{1}{2}\})$. Then $\cup_k M_k$ contains the $1/r$-neighborhood of Γ.

(2) ϕ_k is a homeomorphism of N_k onto B^n and $\phi_k(N_k \cap G) = B_+^n$.

(3) Any $r+1$ distinct N_k's have empty intersection.

(4) ϕ_k and ϕ_k^{-1} are uniformly of class p, and the derivatives of order $\leqq p$ of their components are bounded uniformly with respect to k.

We shall call $\{N_k, \phi_k\}$ a *regular cover* of Γ.

Let $E_+^n = \{x \mid x \in E^n, x_1 > 0\}$. The boundary of E_+^n is $\{x \mid x \in E^n, x_1 = 0\}$, and we identify it with E^{n-1}. Let B^n the open unit sphere in E^n, $B_+^n = B^n \cap E_+^n$, $B^{n-1} = B^n \cap E^{n-1}$.

LEMMA 1.1. *For each non-negative integer q, there is a linear mapping $T_q\colon C_0^\infty(E_+^n \cup E^{n-1}) \to C_c^q(E^n)$ such that $T_q u = u$ on E_+^n and such that for each $0 \leqq s < q+1$, T_q extends to a continuous open mapping of $W^{s,2}(E_+^n)$ into $W^{s,2}(E^n)$.*

Proof. Suppose $u \in C_0^\infty(E_+^n \cup E^{n-1})$. We may suppose that u has already been extended to be continuous on $E_+^n \cup E^{n-1}$. Set $T_q u(x) = u(x)$, $x_1 > 0$, and

$$T_q u(x_1, \cdots, x_n) = \sum_{j=1}^{q+1} \lambda_j u(-jx_1, x_2, \cdots, x_n), \quad x_1 < 0,$$

where the constants λ_j are to be determined. Now $T_q u$ will be in $C_c^q(E^n)$ if

$$\sum_{j=1}^{q+1} (-j)^k \lambda_j = 1, \quad k = 0, 1, \cdots, q.$$

This system can be solved uniquely for the λ_j. It follows from the form of the mapping T_q and the definition of the norms in $W^{s,2}(E_+^n)$ and $W^{s,2}(E^n)$ that T is continuous with respect to them for $0 \leqq s < q+1$. Since $C_c^\infty(E_+^n \cup E^{n-1})$ is dense in each $W^{s,2}(E_+^n)$, T_q has a unique extension with the desired properties. Q. E. D.

LEMMA 1.2. *Let $G \subseteq E^n$ be uniformly regular of class p. Then there are positive constants δ_0 and M, and an integer r, such that for any δ with $0 < \delta < \delta_0$ there is an open cover $\{N_k\}$ of G, a family of homeomorphisms $\{\phi_k\}$ with ϕ_k mapping N_k onto the sphere $B^n(\delta)$ of radius δ about $0 \in E^n$, and a family of functions $\{\eta_k\}$ such that*

(1) *N_k has diameter $\leqq \delta$ and any $r+1$ distinct N_k's have empty intersection;*

(2) $\operatorname{supp}(\eta_k) \subseteq N_k$, $0 \leqq \eta_k(x) \leqq 1$, *and for each $x \in G$, $\Sigma(\eta_k(x))^2 = 1$;*

(3) $\eta_k \in C_0^p(E^n)$ *and* $|D^\alpha \eta_k(x)| \leqq M$, *all* $x, k, |\alpha| \leqq p$;

(4) ϕ_k *and* ϕ_k^{-1} *are uniformly of class* p *and the derivatives of order* $\leqq p$ *of their components are all bounded by* M *in absolute value;*

(5) *If* $N_k \cap \Gamma \neq \emptyset$, *where* Γ *is the boundary of* G, *then* $\phi_k(N_k \cap \Gamma) = B_+^n(\delta) = B^n(\delta) \cap E_+^n$;

(6) *There is a fixed compact set* $K \subseteq B^n(\delta)$ *such that the functions* $\eta_k(\phi_k^{-1}(x))$ *all have support in* K.

This is contained in Lemma 2 of [9]. We shall call the family $\{N_k, \phi_k, \eta_k\}$ a *regular cover* of G.

LEMMA 1.3. *Let* $G \subseteq E^n$ *be uniformly regular of class* p, *and let* $\{N_k, \phi_k, \eta_k\}$ *be a regular cover of* G. *For each* s *such that* $0 \leqq s \leqq p$, *there are positive constants* K_s *and* K_s' *such that for all* $u \in W^{s,2}(G)$, *letting* $[s]$ *be the largest integer* $\leqq s$, *we have*

$$K_s(\| u \|_{s,G})^2 \leqq \Sigma(\| \eta_k^2 u \|_{s,G})^2 \leqq K_s' \sum_{j=0}^{[s]} (\| u \|_{s-j,G})^2.$$

Proof. The cases $s = 0, 1, \cdots, p$ of this lemma are contained in Lemma 3 of [9]. The left-hand inequality is obtained for $q + s$, where $q = 0, 1, \cdots, p-1$ and $0 < s < 1$ by an obvious modification of the argument in [9], using (10). Similarly, in the right-hand inequality we are led to consider terms of the form vw, where $v \in W^{s,2}(G)$ and w and its first derivatives are continuous and bounded by some constant M. We want to show that for some constant N,

$$({}_0\| vw \|_{s,G})^2 \leqq N(\| v \|_{s,G})^2. \tag{14}$$

But set $G_1 = \{(y + x, x) \mid x, y \in G\}$. Now

$$\begin{aligned} | vw(x+y) - vw(x) |^2 \leqq 2 \,| w(x+y) |^2 \,| v(x+y) - v(y) |^2 \\ + 2 \,| v(y) |^2 \,| w(x+y) - w(y) |^2, \end{aligned}$$

so

$$\begin{aligned} ({}_0\| vw \|_{s,G})^2 &= \int\int_{G_1} | vw(x+y) - vw(x) |^2 \,| y |^{-n-2s} \,dxdy \\ &\leqq 2M^2({}_0\| v \|_{s,G})^2 + 2(\| v \|_G)^2 \int_{R^n} \inf(| y |^2, M^2) \,| y |^{-n-2s} \,dxdy, \\ &\leqq N(\| v \|_{s,G})^2. \end{aligned}$$

Q. E. D.

PROPOSITION 1.1. *Let* $G \subseteq E^n$ *be uniformly regular of class* $p \geqq 2$. *There is a linear mapping* $t\colon C_0^\infty(G \cup \Gamma) \to C_0^p(E^n)$ *such that* $Tu = u$ *on*

G and such that for each s with $0 \leqq s \leqq p$, T has a unique extension to a continuous open mapping from $W^{s,2}(G)$ into $W^{s,2}(E^n)$.

Proof. Let $\{N_k, \phi_k, \eta_k\}$ be a regular cover for G; we may assume that $\delta = 1$. Let T_p be the mapping of Lemma 1, extended to $C_\delta^p(E_+^n \cup E^{n-1})$. Given $u \in C_0^\infty(G \cup I)$, let $u_k = \eta_k^2 u$. If $N_k \cap \Gamma$ is empty, let $Tu_k = u_k$. Otherwise define $v_k \in C_\delta^p(E_+^n \cup E^{n-1})$ by putting $v_k = 0$ on $E_+^n - B_+^n$ and $v_k(x) = u_k(\phi_k^{-1}(x))$ on B_+^n. Then set $Tu_k = 0$ on $E^n - N_k$, $Tu_k(x) = (T_p v_k)(\phi_k(x))$, $x \in N_k$. It follows from Lemmas 1, 2, and 3 that T is continuous with respect to the appropriate norms, and since $C_0^\infty(G \cup \Gamma)$ is dense in $W^{s,2}(G)$, the extension of T as a continuous operator is unique for each s. Q. E. D.

COROLLARY. *Let $G \subseteq E^n$ be uniformly regular of class $p \geqq 2$. Suppose $0 \leqq s < t \leqq p$ and j is an integer, $0 \leqq j \leqq p$. Then there are positive constants $C_{s,t}$, $K_{s,t}$, and K_j such that for all u in the appropriate spaces,*

(1) $(\| u \|_{s,G})^2 \leqq C_{s,t}(\| u \|_{t,G})^2$;

(2) $(\| u \|_{s,G})^2 \leqq \epsilon(\| u \|_{t,G})^2 + K_{s,t}\,\epsilon^{-a}(\| u \|_G)^2$,

for all $0 < \epsilon \leqq s/t$, where $a = s/(t-s)$;

(3) $(\| u \|_{j,G})^2 \leqq \sum_{|\alpha|=j} (\| D^\alpha u \|_G)^2 + K_j(\| u \|_G)^2$.

Proof. By Proposition 1 and the equivalence of the norms $\| v \|_s$ and $\| v \|^*_s$, it suffices to prove (1) and (2) for the latter norms. In this case (1) is trivial and (2) follows by taking Fourier transforms and noting that $(1+\lambda)^s - \epsilon(1+\lambda)^t \leqq k_{s,t}\epsilon^{-a}$, all $\lambda \geqq 0$, provided $0 < \epsilon \leqq s/t$ and $k_{s,t} = (s/t)^a(1-s/t)$. Finally, (3) follows from (2), since

$$(\| u \|_{j-1,G})^2 \leqq \tfrac{1}{2}(\| u \|_{j,G})^2 + K(\| u \|_G)^2,$$

so

$$(\| u \|_{j-1,G})^2 \leqq \sum_{|\alpha|=j} (\| D^\alpha u \|_G)^2 + 2K(\| u \|_G)^2,$$

and

$$\begin{aligned}(\| u \|_{j,G})^2 &= \sum_{|\alpha|=j} (\| D^\alpha u \|_G)^2 + (\| u \|_{j-1,G})^2 \\ &\leqq 2\sum_{|\alpha|=j} (\| D^\alpha u \|_G)^2 + 2K(\| u \|_G)^2.\end{aligned}$$

PROPOSITION 1.2. *Let $G \subseteq E^n$ be bounded and uniformly regular of class $p \geqq 2$. If $-p \leqq s < t \leqq p$, then the natural imbedding of $W^{t,2}(G)$ into $W^{s,2}(G)$ is compact.*

This is a well-known consequence of the Sobolev imbedding theorem; see [9, Lemma 5] or [19, Theorem 2.2.3].

Function spaces may also be defined on the boundary Γ of a uniformly regular region G. Suppose G is uniformly regular of class p. Let $\mathcal{D}^p(\Gamma)$ be the space of p-times continuously differentiable functions on Γ, topologized in the usual way. The conjugate space $(\mathcal{D}^p(\Gamma))^*$ is isomorphic to the space of distributions of order p on Γ. Let $\{N_k, \phi_k\}$ be a regular cover of Γ. As before, identify the unit sphere B^{n-1} of E^{n-1} with $\{x \mid x \in B^n, x_1 = 0\}$. Then ϕ_k is a homeomorphism of class p from $N_k \cap \Gamma$ onto B^{n-1}. From the conditions of N_k and ϕ_k it follows as in Lemma 3 that the space

$$W^{j,2}(\Gamma) = \{f \mid f \in (\mathcal{D}^p(\Gamma))^*, D^\alpha f \in L^2(\Gamma) \text{ all } |\alpha| \leqq j\}$$

is the same as the space of all $f \in L^2(\Gamma)$ such that $\| f \|_{j,\Gamma} < \infty$, where

$$(\| f \|_{j,\Gamma})^2 = \sum_k (\| f(\phi_k^{-1}(x)) \|_{j,B^{n-1}})^2. \tag{15}$$

Here j is an integer, $0 \leqq j < p$, and the measure on Γ is that corresponding to the Riemannian metric on Γ induced by the imbedding of Γ in E^n. For any s, $0 \leqq s < p$, we define $W^{s,2}(\Gamma)$ to be the space of all $f \in L^2(\Gamma)$ such that $\| f \|_{s,\Gamma} < \infty$, where

$$(\| f \|_{s,\Gamma})^2 = \sum_k (\| f(\phi_k^{-1}(x)) \|_{s,B^{n-1}})^2. \tag{16}$$

Note that a different regular cover will induce an equivalent norm. Note also that by applying the corollary of Proposition 1 to B^{n-1} and then using (16) we can deduce the analog of that corollary for the spaces $W^{s,2}(\Gamma)$ also.

Denote the L^2-inner product on Γ by $(u, v)_\Gamma$. For $v \in W^{s,2}(\Gamma)$ and $u \in L^2(\Gamma)$, let $\langle u, v \rangle = (u, v)_\Gamma$. This gives an imbedding of $L^2(\Gamma)$ into $(W^{s,2}(\Gamma))^* = W^{-s,2}(\Gamma)$.

If Γ is bounded the cover $\{N_k\}$ is necessarily finite. Therefore from (16) and Proposition 2 we deduce

PROPOSITION 1.3. *Let $G \subseteq E^n$ be uniformly regular of class $p \geqq 2$ and let Γ be its boundary. If Γ is bounded and $-p < s \leqq t < p$, then the natural imbedding of $W^{t,2}(\Gamma)$ into $W^{s,2}(\Gamma)$ is compact.*

The next result establishes a connection between regularity properties of functions on G and their boundary values.

PROPOSITION 1.4. *Let $G \subseteq E^n$ be uniformly regular of class $p \geqq 2$. Let $\partial/\partial n$ be the directional derivative in the direction of the unit inner normal vector at the boundary Γ, and set $D_n = (i)^{-1}\partial/\partial n$. Then the mapping from $C_c^\infty(G \cup \Gamma)$ to $C_c^{p-1}(\Gamma) \oplus \cdots \oplus C_c^{p-1}(\Gamma)$ given by*

$$u \to (u \mid_\Gamma, D_n u \mid_\Gamma, \cdots, D_n^{p-1} u \mid_\Gamma)$$

has a unique extension to a continuous mapping from $W^{p,2}(G)$ *onto* $W^{p-\frac{1}{2},2}(\Gamma) \oplus \cdots \oplus W^{\frac{1}{2},2}(\Gamma)$.

This is proved in [22] (Prop. 8 of Chapter II, § 3) for the case of a region with C^∞ manifold as boundary; there is no difficulty in extending it to the present case.

Given $G \subseteq E^n$, uniformly regular of class p and having boundary Γ, and given $x \in \Gamma$, let $n(x)$ denote the unit inner normal vector to Γ at x. There is an $\epsilon_0 > 0$ such that for $0 \leqq \epsilon \leqq \epsilon_0$ the point set $\Gamma_\epsilon = \{x + \epsilon n(x) \mid x \in \Gamma\}$ is a manifold of class $p-1$, and such that $\pi_\epsilon(x) = x + \epsilon n(x)$ is a homeomorphism of Γ onto Γ_ϵ with π_ϵ and π_ϵ^{-1} uniformly of class $p-1$. Let Γ_ϵ have the measure induced by the natural Riemannian metric. If f is a function defined on Γ_ϵ, let $P_\epsilon f(x) = f(\pi_\epsilon(x))$, $x \in \Gamma$. Then as a mapping from $L^2(\Gamma_\epsilon)$ to $L^2(\Gamma)$, P_ϵ is a topological isomorphism. Furthermore for ϵ_0 small enough, the norms of P_ϵ and P_ϵ^{-1} are bounded uniformly with respect to ϵ.

If u is a measurable function on G, then $u \mid_{\Gamma_\epsilon}$ is measurable for almost all ϵ. Write $P_\epsilon u$ for $P_\epsilon(u \mid_{\Gamma_\epsilon})$. We say that u has *L^2-boundary values* if there is a null set N in the interval $(0, \epsilon_0)$ such that for $0 < \epsilon < \epsilon_0$, $\epsilon \notin N$ we have $P_\epsilon u \in L^2(\Gamma)$ and

$$\lim_{\epsilon, \eta > 0} \| P_\epsilon u - P_\eta u \| = 0.$$

In the following lemma we write E^n as

$$E \times E^{n-1} = \{(t, x) \mid t \in E, x \in E^{n-1}\}.$$

LEMMA 1.4. *There is a positive constant c such that if $u \in W^{1,2}(E^n)$ then for almost all t, s in E,*

$$\int_{E^{n-1}} | u(t,x) - u(s,x) |^2 \, dx \leqq c \mid t - s \mid (\| u \|_1)^2.$$

Proof. Denote the Fourier transforms in E^n and E^{n-1} by $\mathfrak{F}$ and $\mathfrak{F}_x$ respectively. Then

$$\begin{aligned} | \mathfrak{F}_x u(t,\xi) - \mathfrak{F}_x u(s,\xi) |^2 &= c_0 \left| \int (e^{i\tau t} - e^{i\tau s}) \mathfrak{F} u(\tau, \xi) \, d\tau \right|^2 \\ &\leqq c_0 \int \frac{| e^{i\tau t} - e^{i\tau s} |^2}{1 + \tau^2} \, d\tau \int (1 + \tau^2) | \mathfrak{F} u(\tau, \xi) |^2 \, d\tau \\ &\leqq c_1 \mid t - s \mid \int (1 + \tau^2) | \mathfrak{F} u(\tau, \xi) |^2 \, d\tau. \end{aligned}$$

Therefore

$$\begin{aligned} \int_{E^{n-1}} | u(t,x) - u(s,x) |^2 \, dx &\leqq c_1 \mid t - s \mid (\| u \|^*_1)^2 \\ &\leqq c \mid t - s \mid (\| u \|_1)^2. \end{aligned}$$

PROPOSITION 1.5. *Let $G \subseteq E^n$ be uniformly regular of class $p \geqq 2$. Let j be an integer, $1 \leqq j \leqq p$, and suppose $|\alpha| \leqq j-1$. Then $D^\alpha u$ has an L^2-boundary value for all u in $W^{j,2}(G)$.*

Proof. The hypotheses imply that $D^\alpha u \in W^{1,2}(G)$, so it is sufficient to show that each $u \in W^{1,2}(G)$ has L^2-boundary value. Let $\{N_k, \phi_k\}$ be a regular cover of Γ. Modifying ϕ_k we may assume that ϕ_k and ϕ_k^{-1} are uniformly of class $p-1$ with bounds on the derivatives independent of k, while for some positive constant a and all ϵ, $0 < \epsilon < \epsilon_1$, if $x \in N_k \cap \Gamma$ then $\phi(x + \epsilon n(x)) = (a\epsilon, \phi_2(x), \cdots, \phi_n(x))$. Now given $u \in W^{1,2}(G)$, extend it by Proposition 1 to $Tu \in W^{1,2}(E^n)$. Let $v_k(x) = Tu(\phi_k^{-1}(x))$ on B^n. Extending v_k to $W^{1,2}(E^n)$, applying Lemma 4, and translating back into G, we see that $P_\epsilon u \to u|_\Gamma$ in $L^2(\Gamma)$, where $u|_\Gamma$ denotes the extension to $W^{1,2}(G)$ of the corresponding mapping for $u \in C_0^\infty(G \cup \Gamma)$.

Remarks. More details on the spaces $W^{s,2}$ are found in [9], [19], [21], [22]. Lemma 1.1 is found in [21]; the author has been informed that an extension of it similar to Proposition 1.1 has been proved by Aronszajn (unpublished).

Chapter 2. A Priori Estimates for Regular Systems.

Let G be an open set in E^n, $n \geqq 2$, which is uniformly regular of class $2p$ for some positive integer p. Let Γ be the boundary of G. A differential operator of order r defined on G is a mapping $A: \mathcal{D}(G) \to \mathcal{D}^*(G)$ of the form $A = \sum_{|\alpha| \leqq r} a_\alpha(x) D^\alpha$, where $a_\alpha \in \mathcal{D}^*(G)$ and some a_α for $|\alpha| = r$ is not zero. The operator A is said to be *homogeneous* if $a_\alpha = 0$ for $|\alpha| \neq r$. The *principal part* of A is the homogeneous operator $\sum_{|\alpha|=r} a_\alpha D^\alpha$. If the principal part of A has continuous coefficients we define the *characteristic polynomial* $a(x,\xi)$ for $x \in G$ and $\xi \in C^n$, n-dimensional complex space, by

$$a(x,\xi) = \sum_{|\alpha|=r} a_\alpha(x) \xi^\alpha. \tag{1}$$

The operator A is said to be *elliptic* if for each x in G the characteristic polynomial $a(x,\xi)$ has no non-zero roots ξ in R^n. If $n \geqq 3$ it is easily seen that if A is elliptic it must have even order [1, Introduction]. The operator A is said to be *uniformly elliptic* if

$$\inf_{x \in G} \inf_{|\xi|=1, \xi \in R^n} |a(x,\xi)| = c_0 > 0. \tag{2}$$

We shall call c_0 the *constant of ellipticity* of A.

Suppose A is elliptic and of order $2p$, and suppose that the coefficients a_α for $|\alpha| = r$ are defined and continuous on $G \cup \Gamma$. The operator A is said to be ***regularly elliptic*** if for each x in Γ and each non-zero tangent vector t to Γ at x, the polynomial $a(x, t + \lambda n)$ has precisely p roots λ (counting multiplicities) in the upper half plane. Here $n = n(x)$ is the unit inner normal. For $n \geqq 3$ an elliptic operator is necessarily regularly elliptic [1, Introduction].

A differential operator of order r defined at Γ is a mapping

$$B : C_0^\infty(G \cup \Gamma) \to (\mathfrak{D}^{2p}(G \cup \Gamma))^*$$

of the form $B = \sum_{|\beta| \leqq r} b_\beta(x) D^\beta$, where $b_\beta \in (\mathfrak{D}^{2p}(G \cup \Gamma))^*$ and some b_β for $|\beta| = r$ is not zero. The principal part and characteristic polynomial of B are defined as above.

Let $(A, B) = (A, B_1, B_2, \cdots, B_p)$ be a system of operators. We say that this system is ***regular*** if the following conditions are satisfied: (1) The operator A is elliptic and regularly elliptic on G, and is of order $2p$; the operator B_k is of order $r_k < 2p$ and its principal part has continuous coefficients, $1 \leqq k \leqq p$. (2) Furthermore, given $x \in \Gamma$ and t a non-zero tangent vector to Γ at x, let $C = C(x, t)$ be a Jordan curve in the upper half plane containing those roots of $a(x, t + \lambda n)$ having positive imaginary part. Define $p_{kj} = p_{kj}(x, t)$ for $j = 1, 2, \cdots, p$ by

$$p_{kj} = \int_C \lambda^{j-1} b_k(x, t + \lambda n) [a(x, t + \lambda n)]^{-1} d\lambda. \tag{3}$$

Then the matrix $(p_{kj}(x, t))$ is non-singular.

The system (A, B) is said to be ***uniformly regular*** if it is regular and in addition A is uniformly elliptic, while

$$\inf_{x \in \Gamma} \inf_{|t|=1} |\det(p_{kj}(x, t))| = c_1 > 0. \tag{4}$$

We call c_1 the *constant of regularity* of (A, B).

Note that we may also define uniform regularity for any domain G and any open subset Γ_0 of Γ such that Γ_0 is a C^1-manifold.

We can now state the Theorem which is the principal analytical tool of the paper.

THEOREM 2.1. *Let $G \subseteq E^n$ be uniformly regular of class $2p$, with $n \geqq 2$, $p \geqq 1$, and let Γ be the boundary of G. Let A be an elliptic operator of order $2p$ defined on G, having bounded measurable coefficients and with the coefficients of the principal part having bounded uniformly continuous*

first derivatives. For $k=1,2,\cdots,p$ let B_k be an operator of order r_k defined at Γ, having coefficients with bounded distribution derivatives of order $\leqq q_k=2p-1-r_k$, and with the coefficients of the principal part having bounded uniformly continuous first derivatives. Let Γ_ϵ be the surface $\{x+\epsilon n(x)\mid x\in\Gamma\}$, where $n(x)$ is the unit inner normal vector to Γ at x. Then if (A,B) is uniformly regular, there are positive constants N and ϵ_1 such that for any $u\in C_o^\infty(G\cup\Gamma)$;

$$\text{(a)}\quad \| u\|_{2p-\frac{1}{2},G}\leqq N(\| u\|_G+\| Au\|_G+\sum_{k=1}^{p}\| B_k u\|_{q_k,\Gamma}).$$

$$\text{(b)}\quad \sup_{|\alpha|<2p}\sup_{0<\epsilon<\epsilon_1}\| D^\alpha u\|_{\Gamma_\epsilon}\leqq N(\| u\|_G+\| Au\|_G+\sum_{k=1}^{p}\| B_k u\|_{q_k,\Gamma}).$$

Part (a) of this theorem is similar to the estimates obtained in [1], [8], and [22], where $2p-\frac{1}{2}$ is replaced by $2p$ and q_k by $q_k+\frac{1}{2}$. The proof both of parts (a) and (b) is lengthy and technical, and involves no major variations on the technique in [8], hence we do not reproduce it here; details are given in [4]. We note that, in outline, one starts with the assumptions that $G=E_+^n, =\{(t,x)\}$, while A and the B_k are homogeneous operators with constant coefficients. Taking the partial Fourier transform $v=\mathfrak{F}_x u$, we have that for each $\xi\in R^{n-1}$, $v(t,\xi)$ is the solution of a system of ordinary differential equations and can be represented explicitly in terms of the data. This representation provides the estimate in this case. Using the Caldron-Zygmund theorem one proceeds to the case of operators with variable coefficients on a half-space. The general case is then obtained by taking a regular cover of G and mapping each piece into a half-space.

A slight further refinement of this procedure also establishes the following lemma, which will be used in Chapter 4. Let B^n be the unit sphere in E^n, $B^{n-1}=\{x\mid x\in B^n, x_1=0\}$.

LEMMA 2.1. *Let K be a compact subset of $B^{n-1}\subseteq B^n$, and let (c_0,c_1,M) be positive constants. Then there are positive constants c_2 and N such that the following is true: Suppose (A,B) is a uniformly regular system of homogeneous operators on (B^n,B^{n-1}), with constant of ellipticity $\geqq c_0$, constant of regularity $\geqq c_1$, and coefficients bounded by M. Suppose that the functions $a_\alpha(x)-a_\alpha(0)$ and their first derivatives are bounded in absolute value by c_2 and that the functions $b_\beta(x)-b_\beta(0)$ and their derivatives of order $\leqq q_k+1=2p-r_k$ are bounded in absolute value by c_2. Then if $g_k\in C_o^{2p}(B^{n-1})$, $1\leqq k\leqq p$, there is a function $u\in W^{2p,2}(B_+^n)$ such that $Bu=g$ on K, while*

$$\| u \|_{2p-\frac{1}{2}, B_+^n} + \sup_{|\alpha|<2p} \sup_{0<t<\delta} \| D^\alpha u(t,\cdot) \|_{B^{n-1}} + \| Au \|_{B_+^n} \leqq N \sum_{k=1}^{p} \| g_k \|_{q_k, B^{n-1}}.$$

Furthermore the u's can be chosen so that the mapping $g \to u$ is linear.

Chapter 3. Conjugate Boundary Systems.

Let G be an open subset of E^n, and $A = \sum_{|\alpha| \leqq r} a_\alpha D^\alpha$ an operator with coefficients $a_\alpha \in \mathfrak{D}^*(G)$. Then $A : \mathfrak{D}(G) \to \mathfrak{D}^*(G)$. The *formal adjoint* A' is the mapping from $\mathfrak{D}(G)$ to $\mathfrak{D}^*(G)$ given by

$$\langle Au, v \rangle = \langle u, A'v \rangle, \quad u, v \in \mathfrak{D}(G). \tag{1}$$

For $u \in \mathfrak{D}(G)$, $A'u = \sum D^\alpha(\bar{a}_\alpha u)$. Clearly $(A')' = A$. Suppose the coefficients a_α and a_α' of A and A' are summable on bounded sets. Then for any $u, v \in \mathfrak{D}(G)$, Au and Av are in $L^2(G)$ and equation (1) becomes

$$(Au, v)_G = (u, A'v)_G. \tag{2}$$

Here $(u, v)_G$ is the inner product of $L^2(G)$. Note that for $|\alpha| = r$, $a_\alpha' = \bar{a}_\alpha$. Therefore the characteristic polynomial $a'(x, \xi)$ is the complex conjugate of $a(x, \xi)$, and A' is elliptic, uniformly elliptic, or regularly elliptic if and only if A is.

Throughout this chapter we shall assume that G is an open set in E^n with boundary Γ and that G is uniformly regular of class $2p$, while $p \geqq 1$, $n \geqq 2$. Let N_k, $k = 0, 1, \cdots, 2p-1$ be the operator taking $u \to D_n{}^k u \mid_\Gamma$, where D_n is the normal derivative $(i)^{-1}\partial/\partial n$. Expressing N_k as $\sum b_{k\beta} D^\beta$, we see that N_k is of order k, and the $b_{k\beta}$ have bounded continuous derivatives of order $\leqq 2p - k$.

Let $B = (B_0, B_1, \cdots, B_{2p-1})$ be a system of operators defined at Γ, and suppose that B_k is of order k, while $b_{k\beta}$ is continuous for $|\beta| = k$. Then B_k may be written as $\sum_{j \leqq k} b_j(x) N_j + C_k$, where C_k involves only tangential derivatives at $x \in \Gamma$. We shall call the system B a *Cauchy system* if for each x, $b_k(x) \neq 0$. We shall call B a *uniform Cauchy system* if $\inf_k \inf_\Gamma |b_k(x)| > 0$, Equivalently, B is a Cauchy system if for each x, $b_k(x, \lambda n)$ is of degree k in λ, where $n = n(x)$ is the unit inner normal at x; B is a uniform Cauchy system if the coefficient of λ^k in $b_k(x, \lambda n)$ has modulus $\geqq c > 0$ for some c and all x, k.

Suppose A and A' are operators of order $2p$ defined on G and having

locally summable coefficients. Suppose B_k and B_k', $k=0,1,\cdots,2p-1$, are operators of order k defined at Γ, having locally summable coefficients. We say that the system (A,B) and (A',B') are *conjugate* if for all $u,v\in C_0^\infty(G\cup\Gamma)$,

$$(3)\qquad (Au,v)_G-(u,A'v)_G=\sum_{k=0}^{2p-1}(-1)^k(B_ku,B'_{2p-1-k}v)_\Gamma.$$

In particular, if $u,v\in\mathfrak{D}(G)$ then (3) reduces to (2). Therefore if (A,B) and (A',B') are conjugate, A and A' are formal adjoints.

LEMMA 3.1. *For $k=0,1,\cdots,2p-1$, let B_k be an operator of order k defined at Γ, whose coefficients have distribution derivatives of order $\leqq 2p-k$ in $L^\infty(\Gamma)$. Suppose $(B_0,B_1,\cdots,B_{2p-1})$ is a uniform Cauchy system. Then the mapping $u\to(B_0,\cdots,B_{2p-1}u)$, $u\in C_0^\infty(G\cup\Gamma)$, has a unique extension to a continuous mapping from $W^{2p,2}(G)$ onto*

$$W^{2p-\frac12,2}(\Gamma)\oplus\cdots\oplus W^{\frac12,2}(\Gamma).$$

Proof. Each B_k can be written as $\sum_{j\leqq k}C_{jk}N_j$, where C_{jk} is an operator of order $\leqq k-j$ involving only tangential derivatives. Furthermore the tangential distribution derivatives of order $\leqq 2p-k$ of C_{jk} are in $L^\infty(\Gamma)$. By assumption, each $C_{kk}(x)=b_k(x)$, where $|b_k(x)|\geqq c>0$, all $x\in\Gamma$. We may solve inductively for $N_0,N_1,\cdots,N_{2p-1}$ in terms of the B_k's:

$$(4)\qquad \begin{aligned} N_0&=b_0^{-1}B_0\\ N_1&=b_1^{-1}[B_1-C_{10}b_0^{-1}B_0]\\ &\cdots\\ N_k&=\sum_{j\leqq k}D_{jk}B_j.\end{aligned}$$

Here D_{jk} is an operator satisfying the same conditions as C_{jk}. Proposition 1.6 and the conditions satisfied by the C_{jk} imply that B has a unique extension as desired. Solving $Bu=f$ is equivalent to solving $Nu=g$, where $g_k=\sum D_{jk}f_j$. By Proposition 1.6 and the conditions on the D_{jk}, this system can be solved. Therefore the extension of B is onto. Q.E.D.

PROPOSITION 3.1. *Let A be an elliptic operator of order $2p$ defined on G, and let $B=(B_0,B_1,\cdots,B_{2p-1})$ be a Cauchy system. Suppose that a_α has bounded continuous derivatives of order $\leqq|\alpha|$ on $G\cup\Gamma$ and $b_{k\beta}$ has tangential distribution derivatives of order $\leqq 2p-1$ in $L^\infty(\Gamma)$. Then there is a unique system (A',B') such that (A,B) and (A',B') are conjugate. Furthermore, if A is uniformly elliptic and B is a uniform Cauchy system, then the same is true of A' and B'.*

Under stronger regularity assumptions the first part of this proposition is contained in Theorem VI of [2]. (In the terminology of [2], Cauchy systems are called Dirichlet systems.) For completeness we sketch a proof.

Consider first the case $G = E^n_+ = \{(t,x) \mid t > 0, x \in E^{n-1}\}$. As before we identify the boundary Γ with E^{n-1}. Now $A = \sum A_{2p-k} D_t^k$, where A_j, $0 \leqq j \leqq 2p$, is an operator of order $\leqq j$ not involving D_t. Let A_j' be the formal adjoint. Then for $u, v \in C_0^\infty(G \cup \Gamma)$,

$$(A_j u, v)_G = (u, A_j' v)_G. \tag{5}$$

If u, v are in $C_0^1(G \cup \Gamma)$, then

$$(D_t u, v)_G - (u, D_t v)_G = i \int_\Gamma u(0,x) \bar{v}(0,x)\, dx. \tag{6}$$

It follows from (5) and (6) that for any $u, v \in C_0^\infty(G \cup \Gamma)$,

$$\begin{aligned} (Au, v)_G &- (u, A'v)_G \\ &= \sum [(D_t^k u, A_{2p-k}' v)_G - (u, D_t^k A_{2p-k}' v)_G] \\ &= \sum_{k=1}^{2p} \sum_{j=0}^{k-1} [(D_t^{k-1} u, D_t^j A_{2p-k}' v)_G - (D_t^{k-j-1} u, D_t^{j+1} A_{2p-k}' v)_G \\ &= \sum_{k=1}^{2p} i(N_{k-j-1} u, N_j A_{2p-k}' v)_\Gamma \\ &= \sum_{k=0}^{2p-1} (-1)^k (N_k u, N_{2p-1-k}' v)_\Gamma, \end{aligned} \tag{7}$$

where

$$-N_k' = \sum_{j \leqq k} (-1)^k i N_j A_{k-j}'. \tag{8}$$

The systems (A, N) and (A', N') are conjugate. The operator N_k' is of order $\leqq j + (k-j) = k$ and the coefficient of N_k in (9) is $(-1)^k i A_0' = \pm i \bar{a}_0(x)$. Since A is elliptic, $a_0(x) \neq 0$ and therefore N' is a Cauchy system. If (A'', N'') were a second system conjugate to (A, N), we should have $A' = A''$, so for all u, v in $C_0^\infty(G \cup \Gamma)$,

$$\sum (-1)^k (N_{2p-1-k} u, [N_k' - N_k''] v)_\Gamma = 0. \tag{9}$$

Proposition 1.4 implies that $\{Nu \mid u \in C_0^\infty(G \cup \Gamma)\}$ is dense in $L^2(\Gamma) \oplus \cdots \oplus L^2(\Gamma)$, so (9) implies that $N'' = N'$.

Suppose again that $G = E^n_+$ and that B is a Cauchy system satisfying the hypotheses above. As in Lemma 1 we solve the system (4) to get $N_k = \sum D_{jk} B_j$, with D_{jk} of order $\leqq k - j$ and not involving D_t. Replacing N_k in (7) by this expression we can solve as in (8) for a system B'. Uniqueness follows exactly as for N'.

Finally, for a general domain G we can find a regular cover, map each portion at the boundary onto the half-space E^n_+, apply the above, and transfer back to G. The conclusions about uniformity follow readily from the same argument. This completes the outline of proof of Proposition 3.1.

Suppose $(B_r, B_{r+1}, \cdots, B_{r+p-1})$ is a normal system. We shall say that it is of *Dirichlet type* if for each $x \in \Gamma$ and each non-zero tangent vector t, the polynomials $b_{r+k}(x, t+\lambda n)$ are divisible by λ^r for $k = 0, 1, \cdots, p-1$.

PROPOSITION 3.2. *Suppose A is regularly elliptic of order $2p$ with bounded coefficients, and suppose $B = (B_r, B_{r+1}, \cdots, B_{r+p-1})$ is a system of Dirichlet type having bounded coefficients. Then (A, B) is regular. Furthermore, if A is uniformly elliptic and B is uniformly normal, then (A, B) is uniformly regular.*

This was remarked by Browder. By Lemma 1 of [7], regularity is equivalent to the "root-covering condition": for $x \in \Gamma$, t a unit tangent vector at x, n the normal vector at x, let $q(x, t+\lambda n)$ be the product of the p linear factors of $a(x, t+\lambda n)$ having roots in the upper half plane as polynomials in λ. Then the polynomials $b_j(x, t+\lambda n)$ are linearly independent modulo q. This condition is clearly satisfied in the above case, since any linear combination of the b's above is a polynomial or order $\leqq r+p-1$ which is divisible by λ^r, hence can have at most $p-1$ roots in the upper half plane. A suitable modification of this argument, together with the proof of Lemma 1 of [7] yields the conclusion on uniform regularity also.

We conclude this chapter with a result due to Schechter [25, Lemma 2.3]. Suppose A and A' are formally adjoint operators of order $2p$, while $B = (B_{r_1}, B_{r_2}, \cdots, B_{r_p})$ and $B' = (B_{s_1}', B_{s_2}', \cdots, B_{s_p}')$ are normal systems. We shall say that the systems (A, B) and (A', B') are *adjoint* if $s_{p-k-1} = 2p-1-r_k$, $k = 1, 2, \cdots, p$, and if there are Cauchy systems C and C' containing B and B' respectively such that (A, C) and (A', C') are conjugate.

PROPOSITION 3.3. *Suppose (A, B) and (A', B') are conjugate systems. Then (A, B) is regular if and only if (A', B') is.*

Chapter 4. The Operator S.

Throughout this chapter we consider a fixed open set $G \subseteq E^n$, $n \geqq 2$, and fixed systems of operators (A, B) and (A', B'). Let Γ be the boundary of G. We assume that G, Γ, (A, B), and (A', B') satisfy conditions (S1)-(S6) below.

(S1) G is regular of class $4p$ and uniformly regular of class $2p$, where p is an integer $\geqq 1$.

Recall that under this assumption there is a constant $\epsilon_0 > 0$ such that if $0 \leqq \epsilon \leqq \epsilon_0$, then $x \to \pi_\epsilon(x) = x + \epsilon n(x)$ is a homeomorphism of Γ onto the surface Γ_ϵ. Furthermore π_ϵ and π_ϵ^{-1} are uniformly of class $2p$. Let

$$G_\epsilon = G - \bigcup_{0 \leqq \eta \leqq \epsilon} \Gamma_\eta, \qquad K = G \cup \Gamma - G_{\epsilon_0}.$$

(S2) A and A' are operators of order $2p$ whose coefficients are uniformly continuous on $G \cup \Gamma$. Also

$$B = (B_0, B_1, \cdots, B_{2p-1}) \text{ and } B' = (B_0', B_1', \cdots, B_{2p-1})$$

are systems of operators whose coefficients are uniformly continuous on K. For $k = 0, 1, \cdots, 2p-1$, B_k and B_k' are of order k.

For any ϵ with $0 \leqq \epsilon \leqq \epsilon_0$, we consider B and B' as systems defined at Γ_ϵ, the boundary of G_ϵ.

(S3) For each ϵ with $0 \leqq \epsilon \leqq \epsilon_0$, B and B' are uniform Cauchy systems on Γ_ϵ. Furthermore (A, B) and (A', B') are conjugate on $(G_\epsilon, \Gamma_\epsilon)$.

Let Λ_1 be a fixed subset of the set $\{0, 1, \cdots, 2p-1\}$, consisting of precisely p elements. Let Λ_0 be the complementary subset. Let

$$\Lambda_1' = \{k \mid 2p - 1 - k \in \Lambda_0\} \text{ and } \Lambda_0' = \{k \mid 2p - 1 - k \in \Lambda_1\}.$$

We shall assume that the elements of Λ_1 and Λ_1' are in increasing order, while the elements of Λ_0 and Λ_0' are in decreasing order. Let γ_1 be the system of operators $\gamma_1 = (B_k \mid k \in \Lambda_1)$, with the appropriate order, and let γ_1' be defined similarly.

(S4) For each ϵ with $0 \leqq \epsilon \leqq \epsilon_0$, the systems (A, γ_1) and (A', γ_1') are uniformly regular on $(G_\epsilon, \Gamma_\epsilon)$.

(S5) For $|\alpha| = 2p$, the coefficients a_α of A, a_α' of A' have bounded uniformly continuous derivatives of order 1. For $j \in \Lambda_1$ and $k \in \Lambda_1'$, all coefficients of B_j and B_j' have bounded continuous derivatives of order $\leqq 2p - j$ and $2p - k$, respectively, on K.

(S6) The systems (A, γ_1) and (A, γ_1') are *regularizable* in the sense of Browder [7]; that is, there are sequences of systems $(A_k, \gamma_{1,k})$ and $(A_k', \gamma_{1,k}')$ such that for each k the coefficients of all operators are of class $2p$ on $G \cup \Gamma$ or K, such that $(A_k, \gamma_{1,k})$ and $(A_k', \gamma_{1,k}')$ are adjoint, and such that the coefficients converge uniformly on G or K to the corresponding coefficients of the systems (A, γ_1) and (A', γ_1').

Suppose $r \geqq 0$. The space $W_{loc}^{r,2}(G)$ is defined to be the space of all

$u \in \mathfrak{D}^*(G)$ such that $u \mid K_1$ is in $W^{r,2}(K_1)$ for any compact subset $K_1 \subseteq G$. Let Ω_0 be the subset of $W_{loc}{}^{2p-1,2}(G)$ consisting of all those u such that $D^\alpha u$ has an L^2-boundary value for all $|\alpha| \leqq 2p-1$. Proposition 1.5 shows that $W^{2p,2}(G) \subseteq \Omega_0$. If $u \in \Omega_0$ and $|\alpha| \leqq 2p-1$, we let $D^\alpha u\,|_\Gamma$ denote the L^2-boundary value, that is the limit in $L^2(\Gamma)$ of $P_\epsilon(D^\alpha u)$.

LEMMA 4.1. *If $u \in \Omega_0$ and $|\alpha| \leqq 2p-1$, then $D^\alpha u\,|_\Gamma$ is in $W^{k,2}(\Gamma)$, where $k = 2p-1-|\alpha|$.*

Proof. Let $v = D^\alpha u$. Then $v \in W_{loc}{}^{k,2}(G)$ and $D^\beta v$ has L^2-boundary value for $|\beta| \leqq k$. We want to show that $v\,|_\Gamma$ is in $W^{k,2}(\Gamma)$. Let D be a tangential operator of order $\leqq k$, having bounded continuous coefficients. Define C by

$$Cw\,|_{\Gamma_\epsilon} = P_\epsilon^{-1} D P_\epsilon (w\,|_{\Gamma_\epsilon}), \quad 0 < \epsilon < \epsilon_0.$$

Let $N = K - (\Gamma \cup \Gamma_{\epsilon_0})$. Then $C\colon W_{loc}{}^{k,2}(N) \to L_{loc}{}^2(N)$, and C is an operator of order $\leqq k$ having bounded continuous coefficients. Since $u \in \Omega_0$, Cv has an L^2-boundary value. Therefore $DP_\epsilon v = P_\epsilon C v$ converges in $L^2(\Gamma)$. Since also $P_\epsilon v \to v\,|_\Gamma$ in $L^2(\Gamma)$, it follows that $D(v\,|_\Gamma) = \lim (P_\epsilon C v)$. Thus $v\,|_\Gamma \in W^{k,2}(\Gamma)$. Q.E.D.

Letting Λ_0 and Λ_0' be the sets above, we set $\gamma_0 = ((-1)^k B_k \mid k \in \Lambda_0)$ and $\gamma_0' = ((-1)^k B_k \mid k \in \Lambda_0')$. The operators B_k and B_k' may be extended in the obvious way to operators mapping Ω_0 into $W^{2p-1-k,2}(\Gamma)$. We define spaces F_1, F_2 and F_1', F_2' as follows: F_1 is the space $E \oplus E_1 = \{[u,f] \mid u \in E, f \in E_1\}$, where

$$E = L^2(G),$$
$$E_1 = \sum_{k \in \Lambda_0} \oplus\, W^{-k,2}(\Gamma).$$

Let $F_1' = E \oplus E_1'$, where

$$E_1' = \sum_{k \in \Lambda_0'} \oplus\, W^{-k,2}(\Gamma).$$

The spaces E_1 and E_1' each contain the space $H = L^2(G) \oplus \cdots \oplus L^2(G)$. Let the inner products in E and H be denoted by $(u,v)_G$ and $(f,g)_\Gamma$ respectively. Relative to the pairing induced by $(f,g)_\Gamma$ the conjugate spaces of E_1 and E_1' are

$$E_2 = (E_1')^* = \sum_{k \in \Lambda_0'} \oplus\, W^{k,2}(\Gamma),$$
$$E_2' = (E_1)^* = \sum_{k \in \Lambda_0} \oplus\, W^{k,2}(\Gamma).$$

Therefore

$$F_2 = E \oplus E_2 = (F_1')^*$$
$$F_2' = E \oplus E_2' = F_1^*.$$

It follows from Lemma 1 that the extensions of $\gamma_0, \gamma_1, \gamma_0', \gamma_1'$ to Ω_0 satisfy:

$$\gamma_0 u \in H \subseteq E_1;\ \gamma_1 u \in E_2, \tag{1}$$

$$\gamma_0' v \in H \subseteq E_1';\ \gamma_1 v \in E_2'. \tag{2}$$

We shall denote pairings between the conjugate spaces above by $\langle u, v\rangle$, $\langle f, g\rangle$, or $\langle [u,f],[v,g]\rangle = \langle u,v\rangle + \langle f,g\rangle$, as the case may be.

If X and Y are vector spaces and $T: X \to Y$ is a linear operator, we shall denote the domain, range, and null space of T by $D(T)$, $R(T)$, and $N(T)$ respectively.

The *minimal operator* A_m of A is the operator in $L^2(G)$ which is the closure of the restriction of A to $\mathfrak{D}(G)$. The *maximal operator* A_M of A is the operator in $L^2(G)$ which is the restriction of A to the domain $D(A_m) = \{u \in W_{loc}^{2p,2}(G) \cap L^2(G) \mid Au \in L^2(G)\}$. The minimal and maximal operators A_m' and A_M' of A' are defined similarly. Corresponding to the boundary operators γ_1 and γ_1' are the operators A_0 and A_0', which are the restrictions of A_M and A_M' to

$$D(A_0) = \{u \mid u \in W^{2p,2}(G), \gamma_1 u = 0\}$$
$$D(A_0') = \{v \mid v \in W^{2p,2}(G), \gamma_1' v = 0\}.$$

The conditions (S1)-(S6) imply that $(A_m)^* = A_M'$, $(A_m')^* = A_M$, $(A_0)^* = A_0'$, $(A_0')^* = A_0$; see Browder [9; Theorem 5] and [7; Theorem 5].

Let $\Omega = \Omega_0 \cap D(A_M)$, that is Ω is the set of functions $u \in W_{loc}^{2p,2}(G)$ such that u, $Au \in L^2(G)$ and such that $D^\alpha u$ has an L^2-boundary value for all $|\alpha| \leqq 2p-1$. Similarly let $\Omega' = \Omega_0 \cap D(A_M')$. Define operators S: $F_1 \to F_2$ and S': $F_1' \to F_2'$ as follows:

$$D(S) = \{[u, \gamma_0 u] \mid u \in \Omega\};\ S[u, \gamma_0 u] = [Au, \gamma_1 u];$$
$$D(S') = \{[v, \gamma_0' v] \mid v \in \Omega'\};\ S'[v, \gamma_0' v] = [A'v, \gamma_1' v].$$

The systems (A, B) and (A', B') were assumed to be conjugate. It follows that for all $u, v \in C_c^\infty(G \cup \Gamma)$,

$$(Au, v)_G - (u, A'v)_G = (\gamma_0 u, \gamma_1' v)_\Gamma - (\gamma_1 u, \gamma_0' v)_\Gamma, \tag{3}$$

or

$$\langle S[u, \gamma_0 u], [v, \gamma_0' v]\rangle = \langle [u, \gamma_0 u], S'[v, \gamma_0' v]\rangle. \tag{4}$$

Since the operators involved are all continuous from $W^{2p,2}(G)$ to the appropriate spaces, (4) also holds for any u and v in $W^{2p,2}(G)$.

Let S_0 and S_0' be the restrictions of S and S' respectively to the domains

$$D(S_0) = \{[u, \gamma_0 u] \mid u \in C_c^\infty(G \cup \Gamma)\}$$
$$D(S_0') = \{[v, \gamma_0' v] \mid v \in C_c^\infty(G \cup \Gamma)\}.$$

LEMMA 4.2. (a) *$D(S_0)$ and $D(S_0')$ are dense in F_1 and F_1', respectively;*

(b) *$S' \subseteq (S_0)^*$ and $S \subseteq (S_0')^*$;*

(c) *S_0 and S_0' are closable. If S_1 and S_1' are the respective closures, then $S_1 \subseteq S$ and $S_1' \subseteq S'$.*

Proof. (a) Since $C_o^\infty(G \cup \Gamma)$ is dense in $W^{2p,2}(G)$, Lemma 3.1 implies that $\{\gamma_0 u \mid u \in C_o^\infty(G \cup \Gamma)\}$ is dense in H, which is dense in E_1. But also $\mathcal{D}(G)$ is dense in E, so $D(S_0)$ is dense in $E \oplus E_1 = F_1$. Similarly $D(S_0')$ is dense in F'.

(b) Suppose $u \in C_o^\infty(G)$ and $v \in \Omega'$. Let K_1 be a bounded open subset of G whose closure contains $\operatorname{supp}(u)$. Then $v \in W^{2p,2}(K_1 \cap G_\epsilon)$ for any ϵ with $0 < \epsilon \leqq \epsilon_0$. It follows that the analog of (3) holds on G_ϵ:

$$(5) \qquad (Au, v)_{G_\epsilon} - (u, A'v)_{G_\epsilon} = (\gamma_0 u, \gamma_1' u)_{\Gamma_\epsilon} - (\gamma_1 u, \gamma_0' v)_{\Gamma_\epsilon}.$$

As $\epsilon \to 0$, the left side of (5) approaches $(Au, v)_G - (u, A'v)_G$. The operators P_ϵ and P_ϵ^{-1} are uniformly bounded, and there are constants $c(\epsilon)$ such that $c(\epsilon) \to 0$ as $\epsilon \to 0$ and

$$(6) \qquad |(P_\epsilon f, P_\epsilon g)_\Gamma - (f, g)_{\Gamma_\epsilon}| \leqq c(\epsilon) \| f \|_{\Gamma_\epsilon} \| g \|_{\Gamma_\epsilon},$$

for all $f, g \in L^2(\Gamma_\epsilon)$. Therefore the right side of (5) differs from

$$(7) \qquad (P_\epsilon \gamma_0 u, P_\epsilon \gamma_1' v)_\Gamma - (P_\epsilon \gamma_1 u, P_\epsilon \gamma_0' v)_\Gamma$$

by at most

$$(8) \qquad c(\epsilon)\{\| \gamma_0 u \|_{\Gamma_\epsilon} \| \gamma_1' v \|_{\Gamma_\epsilon} + \| \gamma_1 u \|_{\Gamma_\epsilon} \| \gamma_0' v \|_{\Gamma_\epsilon}\}.$$

The conditions on the coefficients of B and B' and the fact that $D^\alpha u$ and $D^\alpha v$ have L^2-boundary values for $|\alpha| \leqq 2p - 1$ imply that the expression (7) approaches

$$(9) \qquad (\gamma_0 u, \gamma_1' v)_\Gamma - (\gamma_1 u, \gamma_0' v)_\Gamma$$

as $\epsilon \to 0$. Since P_ϵ^{-1} is uniformly bounded, the term in braces in (8) is also bounded. Therefore the expression (8) $\to 0$ and equations (3) and (4) hold for the given u and v. This shows that $S' \subseteq (S_0)^*$. Similarly $S \subseteq (S_0')^*$,

(c) Parts (a) and (b) imply that the adjoints of S_0 and S_0' are densely defined. Therefore S_0 and S_0' have closures S_1 and S_1'. Suppose $[u, f]$ is in $D(S_1)$. Then there is a sequence $\{u_m\} \subseteq C_o^\infty(G \cup \Gamma)$ such that $u_m \to u$ in E, $Au_m \to v$ in E, $\gamma_0 u_m \to f$ in E_1, and $\gamma_1 u_m \to g$ in E_2. Since A_M is closed, $u \in D(A_M)$ and $Au = v$. It remains to be shown that $D^\alpha u$ has an L^2-

boundary value for $|\alpha| \leqq 2p-1$. Theorem 2.1 implies that $u_m \to u$ in $W^{2p-\frac{1}{2},2}(G)$. Therefore if $|\alpha| \leqq 2p-1$, $D^\alpha u_m |_{\Gamma_\epsilon} \to D^\alpha u |_{\Gamma_\epsilon}$ in $L^2(\Gamma_\epsilon)$ for almost all ϵ, $0 < \epsilon \leqq \epsilon_0$. Theorem 2.1 also implies that $D^\alpha u_m |_{\Gamma_\epsilon}$ converges in $L^2(\Gamma_\epsilon)$ uniformly with respect to ϵ, $0 < \epsilon \leqq \epsilon_1$. Now

$$\begin{aligned}\| P_\epsilon D^\alpha u - P_\eta D^\alpha u \|_\Gamma \leqq & \| P_\epsilon D^\alpha u - P_\epsilon D^\alpha u_m \|_\Gamma \\ & + \| P_\epsilon D^\alpha u_m - P_\eta D^\alpha u_m \|_\Gamma \\ & + \| P_\eta D^\alpha u_m - P_\eta D^\alpha u \|_\Gamma.\end{aligned}$$

For m large, the first and third terms on the right are small, independent of $\epsilon, \eta \leqq \epsilon_1$. Given m, the middle term goes to zero as $\epsilon, \eta \to 0$. Therefore $P_\epsilon D^\alpha u$ converges in $L^2(\Gamma)$. Thus $u \in \Omega$, and $f = \lim \gamma_0 u_m = \gamma_0 u$, $g = \lim \gamma_1 u_m = \gamma_1 u$. This shows that $S_1 \subseteq S$. Similarly $S_1' \subseteq S'$. Q. E. D.

LEMMA 4.3. *Let X be the space*

$$X = \sum_{k \in \Lambda_0'} \oplus W^{k+\frac{1}{2},2}(\Gamma).$$

There is a continuous linear mapping $L_0 \colon X \to W^{2p,2}(G)$ such that $\gamma_1 L_0 g = g$, all $g \in X$.

Proof. By Lemma 3.1, γ_1 maps $W^{2p,2}(G)$ onto X. Let P be the projection of $W^{2p,2}(G)$ onto the othogonal complement of $N(\gamma_1) \cap W^{2p,2}(G)$. Then $L_0 = P\gamma_1^{-1}$ is the desired mapping. Q. E. D.

Let the norm $\| u \|_S$ be defined for u in Ω by

$$\| u \|_S = \| u \|_E + \| Au \|_E + \| \gamma_1 u \|_{E_2}.$$

Let Ω_1 be the completion of $C_0^\infty(G \cup \Gamma)$ with respect to this norm. Suppose $\{u_m\} \subseteq C_0^\infty(G \cup \Gamma)$ and $\| u_m - u_n \|_S \to 0$. Then $\| u_m - u_n \|_{2p-\frac{1}{2},G} \to 0$ and $\| \gamma_0 u_m - \gamma_0 u_n \|_{E_1} \leqq \| \gamma_0 u_m - \gamma_0 u_n \|_H \to 0$, by Theorem 2.1. It follows that we may consider Ω_1 as contained in Ω, and

$$D(S_1) = \{[u, \gamma_0 u]; u \in \Omega_1\}.$$

We consider Ω_1 as a Banach space with the norm $\| u \|_S$.

LEMMA 4.4. *There is a continuous linear mapping $L \colon E_2 \to \Omega_1$ such that $\gamma_1 L g = g$, all $g \in E_2$.*

Proof. Let $\{N_k, \phi_k, \eta_k\}$ be a regular cover of G. Suppose $N_j \cap \Gamma$ is not empty. Let $T_j u(x) = u(\phi_j(x))$. Define $A^{(j)}$ on $B_+^n(\delta)$ to be the principal part of $T_j^{-1}AT$, and define $\gamma_1^{(j)}$ to be the principal part of $T_j^{-1}\gamma_1 T$. If δ is sufficiently small, each system $(A^{(j)}, \delta_1^{(j)})$ will satisfy the conditions of Lemma 2.1 with respect to a fixed set of constants. For convenience we may assume

that this is true for $\delta = 1$. Let K_1 be a compact subset of B^n such that $\operatorname{supp}(T_j^{-1}\eta_j) \subseteq K_1$, all j.

Suppose we are given $g = (g_k \mid k \in \Lambda_1)$, where each g_k is in $C^{2p}(\Gamma)$. Let $g_{kj} = \eta_j g_k$. By Lemma 2.1 there is a v_j in $W^{2p,2}(B_+^n)$, depending linearly on $g^{(j)} = (g_{kj})$, such that $\delta_1^{(j)} v_j = T_j^{-1} g$ on $K_1 \cap B^{n-1}$ and such that

$$(10) \qquad \| v_j \|_{2p-\frac{1}{2}, B_+^n} + \| A^{(j)} v_j \|_{0, B_+^n} \leqq c \, \| g^{(j)} \|_{E_2}.$$

Let $u_j = \eta_j T_j v_j$. Then $Au_j = \eta_j A T_j v_j + A' v_j = \eta_j T_j A^{(j)} v_j + A'' v_j$, where A' and A'' are operators of order $< 2p$. It follows from this and (10) that

$$(11) \qquad \begin{aligned} \| Au_j \|_{0,G} &\leqq c_1 \| A^{(j)} v_j \|_{0,G} + c_2 \| u_j \|_{2p-1,G} \\ &\leqq c_3 \| g^{(j)} \|_{E_2}. \end{aligned}$$

Similarly,

$$\gamma_1 u_j = \eta_j T_j \delta_1^{(j)} v_j + C^{(j)} u_j = \eta_j g^{(j)} + C^{(j)} u_j = \eta^2{}_j g + C^{(j)} v_j,$$

where $C^{(j)}$ is a system of lower order than γ_1. Let $L_1 g = u = \sum u_j$. Let r be an integer such that any $r+1$ distinct sets N_j have empty intersections. Then (11) implies

$$(12) \qquad \begin{aligned} \| Au \|_{0,G} &\leqq \sum \| Au_j \|_{0,G} \leqq \sum c_3 \| g^{(j)} \|_{E_2} \\ &\leqq r c_3 \| g \|_{E_2}. \end{aligned}$$

Also, $\gamma_1 u = \sum \eta^2{}_j g + \sum C^{(j)} u_j = g + Cu$, where C is a system of lower order. The operator C is bounded, therefore, as an operator from Ω_1 to X.

This shows that L_1 has a continuous extension $L_2 \colon E_2 \to \Omega_1$ such that $\gamma_1 L_2 = I + CL_2$, where I is the identity operator and $C \colon \Omega_1 \to X$ is bounded. Let $L = L_2 - L_0 C L_2$, where L_0 is the operator in Lemma 3. Then $L \colon E_2 \to \Omega_1$ is continuous and $\gamma_1 L = I + CL_2 - CL_2 = I$. Q. E. D.

THEOREM 4.1. *Let G, Γ, A, B, A', B' satisfy conditions* (S1)-(S6) *above. Let the spaces and operators F_1, F_2, F'_1, F'_2 and γ_0, γ_1, γ_0', γ_1' be defined as above. Let Ω (resp. Ω') be the set of all functions u in the domain of the maximal operator A_M (resp. A_M') such that the distribution derivatives of u of order $< 2p$ have L^2-boundary value on Γ. Let $S \colon F_1 \to F_2$ and $S' \colon F_1' \to F_2'$ be defined by $S[u, \gamma_0 u] = [Au, \gamma_1 u]$, u in Ω and $S'[v, \gamma_0' v] = [A'v, \gamma_1' v]$, v in Ω'. Then:*

(a) *S (resp. S') is the closure of the restriction of S (resp. S') to those $[u, \gamma_0 u]$ (resp. $[u, \gamma_0' u]$) for u in $C_0^\infty(G \cup \Gamma)$.*

(b) *S and S' are adjoints.*

Proof. Let S_0, S_0', S_1, S_1', Ω_1 and Ω_1' be defined as above. To prove

both (a) and (b) it is sufficient to show that $S_1^* \subseteq S_1'$ and $(S_1')^* \subseteq S_1$. For S_1 and S_1' are the closures of S_0 and S_0' respectively, so Lemma 2 implies $S_1' \subseteq S' \subseteq (S_1)^* \subseteq S_1'$, and $S_1 \subseteq S \subseteq (S_1')^* \subseteq S_1$. By symmetry we need only show that $(S_1')^* \subseteq S_1$.

Suppose $[v,f] \in D((S_1')^*)$ and $(S_1')^*[v,f] = [w,g]$. Then for u in $C_c^\infty(G \cup \Gamma)$,

$$(A'u,v)_G - (u,w)_G = (\gamma_0'u,g)_\Gamma - (\gamma_1'u,f)_\Gamma. \tag{13}$$

In particular, if $u \in \mathfrak{D}(G)$, then $(A'u,v)_G - (u,w)_G = 0$. Therefore $v \in D(A_M)$ and $Av = w$. Now let L be as in Lemma 4 and set $v_1 = Lg$. For any $u \in D(A_0')$, (13) becomes

$$\begin{aligned}(A'u,v)_G - (u,Av)_G &= (\gamma_0'u,g)_\Gamma \\ &= (\gamma_0'u,\gamma_1 v_1)_\Gamma = (A'u,v_1)_G - (u,Av_1)_G.\end{aligned} \tag{14}$$

Therefore $v - v_1 \in D((A_0')^*) = D(A_0) \subseteq W^{2p,2}(G) \subseteq \Omega_1$. But $v_1 = Tg \in \Omega_1$, so $v = (v - v_1) + v_1 \in \Omega_1$. Thus for all $u \in W^{2p,2}(G)$,

$$\begin{aligned}(\gamma_0'u,g)_\Gamma - (\gamma_1'u,f)_\Gamma &= (A'u,v)_G - (u,Av)_G \\ &= (\gamma_0'u,\gamma_1 v)_\Gamma - (\gamma_1'u,\gamma_0 v)_\Gamma.\end{aligned} \tag{15}$$

Now (γ_0',γ_1') maps $W^{2p,2}(G)$ onto a dense subset of $E_1' \oplus E_2'$, so (15) implies that $g = \gamma_1 v$, $f = \gamma_0 v$. Thus $[v,f] = [v,\gamma_0 v] \in D(S_1)$, and $(S_1')^* \subseteq S_1$. This completes the proof of Theorem 1.

COROLLARY. (a) $W^{2p,2}(G) \subseteq \Omega \subseteq W^{2p-\frac{1}{2},2}(G)$;

(b) *Ω depends only on the principal part of A;*

(c) $\Omega' = \{v ; \bar{v} \in \Omega\}$.

Proof. Part (a) follows from Theorem 1, since the inclusions are true for Ω_1 and by the Theorem $\Omega = \Omega_1$.

Part (b) follows from (a), for suppose C is of order $< 2p$ and has bounded coefficients. Then $u \in \Omega$ implies $u \in W^{2p-\frac{1}{2},2}(\Omega)$, so $Cu \in L^2(G)$. Therefore $(A + C)u \in L^2(G)$ and $u \in D((A+C)_M)$.

Part (c) follows from (b), for let $\tilde{A}$ and $\tilde{A}'$ be the principal part of A and A' respectively. Then $\tilde{A}'$ is the complex conjugate of $\tilde{A}$, so $\tilde{A}'v \in L^2(G)$ if and only if $\tilde{A}\bar{v} \in L^2(G)$. Now $A - \tilde{A}$ and $A - \tilde{A}'$ are of lower order, so (b) implies (c).

Chapter 5. Non-Local Boundary Value Problems.

Let A be an elliptic operator defined on an open set G, and let A_m, A_M be the minimal and maximal operators. A *realization* of A is an operator A_1 acting in $L^2(G)$ such that $A_m \subseteq A_1 \subseteq A_M$. A realization may be considered as an abstract boundary value problem; see Vishik [27]. In particular the realization A_0 in Chapter 4 corresponded to the boundary conditions $\gamma_1 u = 0$. These are local boundary conditions, and such problems are amenable to study by the techniques of Chapter 2; see [1], [7], [8], [22], [24], [25].

The point of introducing the operator S in Chapter 4 is to enable one to study realizations of A corresponding to more general boundary conditions; see [14], [3], [16], [17] where this is done for the Laplace operator and general formally self-adjoint second order operators. Suppose A, B, A', B', G and Γ are as in Chapter 4. Suppose $C: \Omega \to E_2$ is a linear operator. Define $C_1: F_1 \to F_2$ by $C_1[u, \gamma_0 u] = [0, Cu]$. The problem is to derive the properties of the realization of A corresponding to the conditions $\delta_1 u = Cu$ from the properties of the perturbed operator $S - C_1$. This reverses the procedure of Chapter 4, where we used properties of A_0, corresponding to $C \equiv 0$, to study S. In particular we shall be interested in when the realization is closed, what its adjoint is, and what its domain is. We begin with an abstract discussion of the problem.

If E and F are Banach spaces and $T: E \to F$ is a linear operator, define the norm $\| u \|_T$ on $D(T)$ by $\| u \|_T = \| u \| + \| Tu \|$. Then $D(T)$ is complete with respect to this norm if and only if T is closed. Let $C: E \to F$ be a second linear operator. We say that C is *T-compact* if $D(C) \supseteq D(T)$ and C is compact as a mapping from $D(T)$ with norm $\| u \|_T$ into F. Similarly we say that C is *T-bounded* if $D(C) \supseteq D(T)$ and C is bounded from $D(T)$ to F. We say that C is *(T, ϵ)-bounded* if C is bounded by ϵ as a mapping from $D(T)$ to F.

Let E, E_1, and E_2 be reflexive Banach spaces, with conjugate spaces $E' = E^*$, $E_2' = E_1^*$, and $E_1' = E_2^*$. We denote all norms simply by $\| \ \|$ and all pairings between conjugate spaces by $\langle \ , \ \rangle$. Let

$$F_1 = E \oplus E_1, \quad F_2 = E \oplus E_2,$$
$$F_1' = E' \oplus E_1', \quad F_2' = E' \oplus E_2'.$$

Let A, δ_0, δ_1, A', δ_0', and δ_1' be densely defined operators,

$$A: E \to E, \quad \delta_0: F \to E_1, \quad \delta_1: E \to E_2$$
$$A': E' \to E', \quad \delta_0': F' \to E_1', \quad \delta_1': E' \to E_2'.$$

We suppose that there are dense subspaces $\Omega \subseteq E$ and $\Omega' \subseteq E'$ such that

$$\Omega \subseteq D(A) \cap D(\delta_0) \cap D(\delta_1)$$
$$\Omega' \subseteq D(A') \cap D(\delta_0') \cap D(\delta_1').$$

Let $T: F_1 \to F_2$ and $T': F' \to F_2'$ be defined by

$$D(T) = \{[u, \delta_0 u] \mid u \in \Omega\}, \quad T[u, \delta_0 u] = [Au, \delta_1 u];$$
$$D(T') = \{[v, \delta_0' v] \mid v \in \Omega'\}, \quad T'[v, \delta_0' v] = [A'v, \delta_1' v].$$

We define conditions (T1), (T2), $\cdots$, as follows.

(T1) $N(\delta_1) \cap \Omega$ is dense in E.

(T2) T is closed.

(T3) There is a constant M such that for all $u \in \Omega$,

$$\| \delta_0 u \| \leqq M(\| u \| + \| Au \| + \| \delta_1 u \|).$$

For $u \in \Omega$, let $\| u \|_T = \| u \| + \| Au \| + \| \delta_1 u \|$. Then conditions (T2) and (T3) imply that Ω is a Banach space with respect to this norm. Furthermore A, δ_0, and δ_1 are all continuous with respect to this norm.

(T4) $\delta_1(\Omega)$ is closed in E_2 and there is a continuous linear mapping $L: \delta_1(\Omega) \to \Omega$ such that $\delta_1 L$ is the identity on $\delta_1(\Omega)$.

Note that if E, and E_2 are Hilbert spaces then (T4) is satisfied whenever $\delta_1(\Omega)$ is closed in E_2. For let P be the projection of Ω onto the complement of $N(\delta_1) \cap \Omega$. Then we can take $L = P\delta_1^{-1}$.

(T5) $T^* = T'$.

Let (T1)$'$, (T2)$'$, $\cdots$, (T5)$'$ be the analogous conditions for T'.

Let A_0 be the restriction of A to $D(A_0) = \{u; u \in \Omega, \delta_1 u = 0\}$ and A_0' the restriction of A' to $D(A_0') = \{v \mid v \in \Omega', \delta_1' v = 0\}$.

THEOREM 5.1. (a) *If conditions* (T1)-(T3) *are satisfied, then* A_0 *is densely defined and closed.*

(b) *If conditions* (T1)-(T5) *and* (T1)$'$-(T5)$'$ *are satisfied, then* $A_0^* = A_0'$ *and* $(A_0')^* = A_0$.

Proof. (a) By (T1), A_0 is densely defined. If T is closed and (T3) holds, clearly A_0 is closed.

(b) Suppose $v \in D(A_0^*)$, $A_0^* v = w$. The mapping $u \to \langle Au, v\rangle - \langle u, w\rangle$ then depends only on $\delta_1 u$, for $u \in \Omega$. Therefore we may consider it as a mapping ϕ from $\delta_1(\Omega)$ to $\boldsymbol{C}$. Now $\delta_1 u = \delta_1 L \delta_1 u$, and

$$| \phi(\delta_1 u) | = | \langle AL\delta_1 u, v\rangle - \langle L\delta_1 u, w\rangle | \leqq M \| L\delta_1 u \|_T$$
$$\leqq M' \| \delta_1 u \|.$$

Therefore $\phi \in (\delta_1(\Omega))^*$. By the Hahn-Banach theorem there is a $g \in E_2^* = E_1'$ such that for all $u \in \Omega$,

$$\langle Au, v\rangle - \langle u, w\rangle = \phi(\delta_1 u) = -\langle \delta_1 u, g\rangle.$$

Therefore $[v, g] \in D(T^*) = D(T')$ and $[w, 0] = T'[v, g] = [A'v, \delta_1' g]$. Thus $v \in D(A_0')$. This shows that $A_0^* \subseteq A_0'$. The reverse inclusion follows from (T5), so $A_0^* = A_0'$. By symmetry, $(A_0')^* = A_0$. Q.E.D.

Now we wish to find conditions on perturbations of T which will preserve (T1)-(T3), and conditions on perturbations of T and T' which will preserve (T1)-(T5) and (T1)$'$-(T5)$'$. The perturbations will be of the form C: $F_1 \to F_2$, $R(C) \subseteq (0) \oplus E_2$. Then

$$N(C^*) \supseteq ((0) \oplus E_2)^\perp = E' \oplus (0).$$

Now $(T - C)^* \supseteq T^* - C^*$, so presumably if C' is to be the corresponding perturbation of T' we should have $C' = C^*$, so $N(C') \supseteq E' \oplus (0)$. Similarly we should expect that $N(C) \supseteq E \oplus (0)$. For this reason we shall consider only operators $C_1 \colon F_1 \to F_2$ which are of the form $C_1[u, f] = [0, Cf]$, where $C \colon E_1 \to E_2$. Given such an operator C, we call C_1 the *induced operator.* With operators of this form, (T1) and (T1)$'$ will be implied by

(T6) $\Omega \cap N(\delta_0) \cap N(\delta_1)$ is dense in E,

(T6)$'$ $\Omega' \cap N(\delta_0') \cap N(\delta_1')$ is dense in E'.

To study the remaining conditions, we begin with two simple lemmas.

LEMMA 5.1. *Let X and Y be Banach spaces and let T and C be linear operators from X to Y, with $D(C) \supseteq D(T)$.*

(a) *Suppose there are positive constants ϵ and K, with $\epsilon < 1$, such that for all $x \in D(T)$,*

$$(*) \qquad \| Cx \| \leq \epsilon \| Tx \| + K \| x \|.$$

Then the norms $\| x \|_T$ and $\| x \|_{T-C}$ are equivalent.

(b) *If C is closable and T-compact, then for any $\epsilon > 0$ there is a constant $K = K(\epsilon)$ such that $(*)$ holds.*

Proof. (a) For $x \in D(T)$,

$$\begin{aligned} \| x \|_{T-C} &= \| x \| + \| (T - C)x \| \leq \| x \| + \| Tx \| + \| Cx \| \\ &\leq (1 + \epsilon + K) \| x \|_T; \\ \| x \|_T &= \| x \| + \| Tx \| \leq \| x \| + \| (T - C)x \| + \| Cx \| \\ &\leq \| x \|_{T-C} + \epsilon \| Tx \| + K \| x \|, \end{aligned}$$

or

$$\| x \|_T \leq (1 - \epsilon)^{-1}(1 + K) \| x \|_{T-C}.$$

(b) Given $\epsilon \geqq 0$, suppose (*) does not hold identically for any $K > 0$. We may assume $\epsilon < 1$. There is a sequence $\{x_n\} \subseteq D(T)$ such that $\| x_n \| = 1$, $\| Cx_n \| \geqq \epsilon \| Tx_n \| + n = \delta_n$. Let $x_n' = \delta_n^{-1} x_n$. Then $\| x_n' \|_T \leqq \epsilon^{-1} + 1$. Since C is T-compact, by choosing a subsequence if necessary we may suppose that $Cx_n' \to y \in Y$. Then $\| Cx_n' \| = \delta_n^{-1} \| Cx_n \| \geqq 1$, so $\| y \| \geqq 1$. However $\| x_n' \| \leqq 1/n$, so $x_n' \to 0$. This contradicts the assumption that C is closable.

LEMMA 5.2. *Let X and Y be Banach spaces and let T be a continuous linear mapping of X onto Y. Let $S: Y \to X$ be a continuous linear mapping such that $TS = I$, the identity operator in Y. Let $C: X \to Y$ be linear. If C is compact or has norm $\| C \| < \| S \|^{-1}$, then $T - C$ has closed range and there is a continuous linear mapping $S_1: R(T-C) \to X$ such that $(T-C)S_1 = I$ on $R(T-C)$.*

Proof. $R(T-C) = R((T-C)S) = R(I-CS)$. If $\| C \| < \| S \|^{-1}$, then $\| CS \| < 1$. In this case $R(I-CS) = Y$ and $(I-CS)^{-1} = \sum_{n=0}^{\infty} (CS)^n$ exists. Let $S_1 = S(I-CS)^{-1}$.

If C is compact, so is CS. In this case the Riesz theory of compact operators [23] implies that $N(I-CS)$ has finite dimension and $R(I-CS)$ is closed. There is a bounded projection $P: X \to N(I-CS)$, and we may take $S_1 = S(I-P)(I-CS)^{-1}$ on $R(I-CS)$.

To apply Lemma 2 we shall want the conditions

(T7) $\delta_1(\Omega) = E_2$

(T7)′ $\delta_1'(\Omega) = E_2'$.

THEOREM 5.2. *Suppose T and T' satisfy conditions* (T2)-(T7) *and* (T2)′-(T7)′. *Let $C: E_1 \to E_2$ have domain $D(C) \supseteq \delta(\Omega)$, and let $C_1: F_1 \to F_2$ be the induced operator. Let C^* be the adjoint of C.*

(a) *Suppose there are positive constants ϵ and K, with $\epsilon < 1$, such that for all $u \in \Omega$,*

$$\| C\delta_0 u \| \leqq \epsilon \| T[u, \delta_0 u] \| + K \| u \|.$$

Then the operators $T - C_1$, $\delta_1 - C\delta_0$ satisfy (T1)-(T3).

(b) *Suppose $D(C^*) \supseteq \delta_0'(\Omega')$. Let $C_1': F_1' \to F_2'$ be the operator induced by C^*. Suppose $(T - C_1)^* = T' - C_1'$. There are constants $\eta > 0$, $\eta' > 0$ such that if C_1 is either T-compact or (T, η)-bounded and C_1' is either T'-compact or (T', η')-bounded, then $T - C_1$ and $T' - C_1'$ satisfy* (T1)-(T5) *and* (T1)′-(T5)′.

Proof. (a) Since (T6) is assumed, (T1) is immediate for $\delta - C\delta_0$.

Suppose the inequality holds. Since $C_1[u, \delta_0 u] = [0, C\delta_0]$, Lemma 1 applies to C_1 and T. Therefore the $(T - C_1)$-norm is equivalent to the T-norm on $D(T)$. Since T is closed, $T - C_1$ is closed. Finally, an argument similar to that of Lemma 1(a) shows that the T-norm on Ω is equivalent to the $(T - C_1)$-norm on Ω, so (T3) holds also.

(b) The assumptions imply that $D(C_1^*)$ is dense in F_1', so C_1 is closable. Suppose C_1 is T-compact. By (T3), $\{[u_n, \gamma_0 u_n]\}$ is bounded in T-norm if and only if $\{\| u_n \|_T\}$ is bounded. Therefore $u \to [0, C\delta_0 u]$ is compact with respect to $u \to [Au, \delta_0 u]$, and the inequality in (a) is satisfied for any $\epsilon > 0$ and suitable $K = K(\epsilon)$. On the other hand, if C_1 is (T, η) bounded with $\eta \leqq \epsilon$ and $M\eta \leqq \epsilon$, then (T3) implies that the inequality in (a) is satisfied with $K = \epsilon$. In either case, (T1)-(T3) hold for $T - C_1$. By symmetry, (T1)$'$-(T3)$'$ hold for $T' - C_1'$. We are assuming (T5), and (T5)$'$ follows from (T2) and (T5). Therefore we need only show that (T4) and (T4)$'$ hold for $\delta_1 - C\delta_0$ and $\delta_1' - C^*\delta_0'$. Now δ_1 is continuous from Ω to E_2, and by (T7) it is onto. Consequently if C_1 is either T-compact or (T, η)-bounded for small enough η, the hypotheses of Lemma 3 will be satisfied for δ_1 and $C\delta_0$ as mappings from Ω to E_2. Thus (T4) will hold for $\delta_1 - C\delta_0$. By symmetry (T4)$'$ will hold for $\delta_1' - C^*\delta_0'$. This completes the proof of Theorem 2.

For the remainder of this chapter we assume that G, Γ, A, B, A', and B' satisfy conditions (S1)-(S6) of Chapter 4. The spaces E, E_1, E_2, F_1, F_2, Ω, E', E_1', E_2', F_1', F_2', and Ω' and the operators γ_0, γ_1, S, γ_0', γ_1' and S' are as defined in Chapter 4. We denote the norms and inner products simply by $\| \ \|$, $\langle \ , \ \rangle$.

THEOREM 5.3. *Let G, Γ, A, B, A', B' satisfy conditions* (S1)-(S6) *of Chapter* 4, *and let the spaces Ω, E_1, E_2 and operators γ_0, γ_1 be as defined in Chapter* 4. *Let $C\colon E_1 \to E_2$ be such that $D(C) \supseteq \gamma_0(\Omega)$. Let $A(C)$ be the restriction of A to those u in Ω such that $\gamma_1 u = C\gamma_0 u$. Suppose there are positive constants $\epsilon < 1$ and K such that for all $u \in \Omega$,*

$$\| C\gamma_0 u \| \leqq \epsilon \| S[u, \gamma_0 u] \| + K \| u \|.$$

Then $A(C)$ is a densely defined closed realization of A.

Proof. This will follows immediately from Theorems 1 and 2, provided we show that S and S' satisfy conditions (T2)-(T7) and (T2)$'$-(T7)$'$ above. By symmetry we need only show that S satisfies (T2)-(T7). As remarked earlier, (T2) is a consequence of Theorem 4.1. Conditions (T4) and (T7) are precisely Lemma 4.4. Condition (T5) is Theorem 4.1(b). Condition

(T6) follows from the fact that $\mathfrak{D}(G)$ is dense in E. This completes the proof of Theorem 3.

Let X and X' be the spaces

$$X = \sum_{k \in \Lambda_0'} W^{k+\frac{1}{2},2}(\Gamma), \quad X' = \sum_{k \in \Lambda_0} W^{k+\frac{1}{2},2}(\Gamma).$$

THEOREM 5.4. *Let G, Γ, A, B, A', B' satisfy conditions* (S1)-(S6) *of Chapter 4, and let the associated spaces and operators be as defined there. Let X and X' be as above. Let $C: E_1 \to E_2$ and $C': E_1' \to E_2'$ be such that $D(C) \subseteq \gamma_0(\Omega)$, $D(C') \subseteq \gamma_0'(\Omega')$, $R(C) \subseteq X$, $R(C') \subseteq X'$. Let $A(C)$ and $A'(C')$ be defined as in Theorem* 5.3. *Suppose C and C' satisfy the condition of Theorem* 5.3. *Then:*

(a) *$A(C)$ and $A'(C')$ are closed and have domains included in $W^{2p,2}(G)$.*

(b) *Let C_1 and C_1' be the operators induced by C and C'. Suppose $(S - C_1)^* = S' - C_1'$. Then $A(C)$ and $A'(C')$ are adjoints.*

Proof. (a) The fact that $A(C)$ and $A'(C')$ are closed follows from Theorem 3. Suppose $u \in D(A(C))$. Then $\gamma_1 u = C\gamma_0 u \in X$. By Lemma 3.1, there is a $v \in W^{2p,2}(G)$ such that $\gamma_1 v = C\gamma_0 u$. Then $u - v \in D(A(0)) \subseteq W^{2p,2}(G)$, so $u = (u - v) + v \in W^{2p,2}(G)$. By symmetry, $D(A'(C')) \subseteq W^{2p,2}(G)$ also.

(b) We want to show that (T1)-(T5) and (T1)$'$-(T5)$'$ hold for $S - C_1$ and $S' - C_1'$. Of these, only (T4) and (T4)$'$ need demonstration. As remarked after (T4) above, we need only show that $(\gamma_1 - C\gamma_0)(\Omega)$ and $(\gamma_1' - C'\gamma_0')(\Omega')$ are closed. Suppose $g \in E_2$. There is a $u = Lg \in \Omega$ such that $\nu_1 u = g$. By Lemma 3.1 there is a $v \in W^{2p,2}(G)$ such that $\gamma_1 v = C\gamma_0 u \in X$ and $\gamma_0 v = 0$. Then

$$(\gamma_1 - C\gamma_0)(u + v) = \gamma_1 u - C\gamma_0 u + C\gamma_0 u = \gamma_1 u = g.$$

Therefore $(\gamma_1 - C\gamma_0)(\Omega) = E_2$. Similarly $(\gamma_1' - C'\gamma_0)(\Omega') = E_2'$. This completes the proof of Theorem 4.

As a straightforward application of Theorems 1, 2, and 3 we have the following:

THEOREM 5.5. *Let $C: E_1 \to E_2$ and $C': E_1' \to E_2'$ be such that $D(C) \supseteq \gamma_0(\Omega)$ and $D(C') \supseteq \gamma_0'(\Omega')$. Let C_1 and C_1' be the induced operators, and assume $(S - C_1)^* = S' - C_1'$. There are positive constants η and η' such that if C_1 is either S-compact or (S, η)-bounded and C_1' is either S'-compact or (S', η')-bounded, then $A(C)$ and $A'(C')$ are adjoints.*

Clearly $\gamma_0(\Omega) \subseteq \sum_{k \in \Lambda_0} \oplus W^{2p-1-k,2}(\Gamma)$. Suppose C_{jk}, $j \in \Lambda_0'$, $k \in \Lambda_0$, is an operator such that

(T8) $C_{jk}: W^{-k,2}(\Gamma) \to W^{j,2}(\Gamma)$,

(T9) $D(C_{jk}) \supseteq W^{2p-1-k,2}(\Gamma)$.

Let $C: E_1 \to E_2$ be defined by the matrix (C_{jk}), that is $(Cf)_j = \sum_k C_{jk} f_k$. Then (T9) implies that $D(C) \supseteq \gamma_0(\Omega)$. The analogous conditions for C_{kj}', $j \in \Lambda_0'$, $k \in \Lambda_0$, are that

(T8)′ $C_{jk}': W^{-j,2}(\Gamma) \to W^{k,2}(\Gamma)$,

(T9)′ $D(C_{kj}') \supseteq W^{2p-1-j,2}(\Gamma)$.

The condition that $C' = C^*$ is that $C_{kj}' = C_{jk}^*$, all j, k. Conditions on the C_{jk} and C_{kj}' that C_1 and C_1' be S-compact and S'-compact are

(T10) C_{jk} is compact from $W^{2p-1-k,2}(\Gamma)$ to $W^{j,2}(\Gamma)$,

(T10)′ C_{kj}' is compact from $W^{2p-1-j,2}(\Gamma)$ to $W^{k,2}(\Gamma)$.

Theorems 2, 3, and 4 have obvious restatements in terms of the operators C_{jk} and C_{kj}'.

As a special case, suppose $\Lambda_1 = (p, p+1, \cdots, 2p-1)$. Then also $\Lambda_1' = (p, p+1, \cdots, 2p-1)$, and for $j \in \Lambda_0'$, $k \in \Lambda_0$ we have $2p-1-k \geqq p > j$.

THEOREM 5.6. *Suppose* $\Lambda_1 = (p, p+1, \cdots, 2p-1)$. *Suppose that for each* j, k, $0 \leqq j, k \leqq p-1$, C_{jk} *and* C_{kj}' *are bounded operators in* $W^{p-\frac{1}{2},2}(\Gamma)$. *Let* $C: E_1 \to E_2$ *and* $C': E_1' \to E_2'$ *be determined by the matrices* (C_{jk}) *and* (C_{kj}'). *Let* C_1 *and* C_1' *be the induced operators.*

(a) $A(C)$ *and* $A'(C')$ *are closed and the domains are contained in* $W^{2p,2}(G)$.

(b) *Suppose* $(S - C_1)^* = S' - C_1'$. *Then* $A(C)$ *and* $A'(C')$ *are adjoints.*

Proof. Clearly under these conditions, γ_0 is continuous from $W^{2p-1,2}(G)$ to $W^{p-\frac{1}{2},2}(G) \oplus \cdots \oplus W^{p-\frac{1}{2},2}(G)$. Therefore $C\gamma_0$ is continuous from $W^{2p-1,2}(G)$ to X. By the corollary to Proposition 1.1, for every $\epsilon > 0$ there is a $K = K(\epsilon)$ such that

$$\| u \|_{2p-1,G} \leqq \epsilon \| u \|_{2p-\frac{1}{2},G} + K \| u \|_{0,G}.$$

The imbedding $\Omega \to W^{2p-\frac{1}{2},2}(G)$ is continuous. Therefore an inequality of the form in Theorem 3 holds for $C\delta_0$. By symmetry this is also true for $C'\delta_0'$. Theorem 6 now follows immediately from Theorem 4.

This theorem obviously may be generalized in various ways, using Theorem 4 and the inequalities in the corollary to Proposition 1.1.

Chapter 6. Operators on Bounded Domains.

In the case of a bounded domain G, if $t > s$ the imbeddings $W^{t,2}(G) \to W^{s,2}(G)$ and $W^{t,2}(\Gamma) \to W^{s,2}(\Gamma)$ are compact. This fact makes it possible to strengthen the results of Chapter 6. We begin a discussion of Fredholm operators; cf. [10, Lemma 1], [18, Theorem 2.6], [5].

LEMMA 6.1. *Let X and Y be Banach spaces, and let $T: X \to Y$ be a closed linear operator. Suppose that the injection of $D(T)$ into X is compact, relative to the norm $\| x \|_T$ on $D(T)$. Then $N(T)$ is finite dimensional and $R(T)$ is closed.*

Proof. The norms $\| x \|_T$ and $\| x \|$ are the same on $N(T)$, so the unit sphere of $N(T)$ is pre-compact. Therefore $N(T)$ is finite dimensional.

Considered as an operator from $D(T)$ to Y, T is continuous. Therefore to prove that $R(T)$ is closed it is sufficient to show that T maps closed bounded sets in $D(T)$ onto closed sets in Y; see [15, Theorem VI. 6. 5]. If $B \subseteq D(T)$ is closed and bounded and y is in the closure of $T(B)$, then there is a sequence $\{x_n\} \subseteq B$ such that $Tx_n \to y$. Since $D(T) \to X$ is compact, by choosing a subsequence if necessary we may assume that $x_n \to x$ in X. Then $x \in B$, and since T is closed, $Tx = y$. Q.E.D.

Let X and Y be Banach spaces. A closed densely defined linear operator $T: X \to Y$ is said to be a *Fredholm operator* if $N(T)$ has finite dimension, $R(T)$ is closed, and $R(T)$ has finite co-dimension in Y. The *index* $\operatorname{ind}(T)$ is

$$(1) \qquad \operatorname{ind}(T) = \dim N(T) - \operatorname{codim} R(T).$$

(We shall also use formula (1), taking the closure of $R(T)$ on the right, to define the index for any operator for which either $\dim N(T)$ or $\operatorname{codim} \bar{R}(T)$ is finite.) Let $\Phi(X, Y)$ denote the collection of all Fredholm operators mapping X into Y. For convenience we shall assume in what follows that X and Y are reflexive.

LEMMA 6.2. *Suppose $T: X \to Y$ is densely defined and closed. Then $T \in \Phi(X, Y)$ if and only if $T^* \in \Phi(Y^*, X^*)$. If so, then* $\operatorname{ind}(T^*) = -\operatorname{ind}(T)$.

LEMMA 6.3. *If $S, T \in \Phi(X, Y)$ and $S \subseteq T$, then* $\operatorname{ind}(S) \leq \operatorname{ind}(T)$. *Equality holds if and only if $S = T$.*

LEMMA 6.4. *Suppose $T \in \Phi(X, Y)$.*

(a) *If $C: X \to Y$ is linear, closable, and T-compact, then $T - C \in \Phi(X, Y)$ and* $\operatorname{ind}(T - C) = \operatorname{ind}(T)$.

(b) *There is an* $\epsilon > 0$ *such that if* $C: X - Y$ *is linear and* (T,ϵ)*-bounded, then* $T - C \in \Phi(X,Y)$ *and* $\operatorname{ind}(T-C) = \operatorname{ind}(T)$.

(c) *If* $C: X \to Y$ *is* T*-compact and* $C^*: X^* \to Y^*$ *is* T^**-compact, then* $(T-C)^* = T^* - C^*$.

Proof of Lemma 6.2. Since X is reflexive and T is closed, T^* is densely defined. The range of T is closed if and only if the range of T^* is; see [10, Theorem A]. In this case $\dim N(T) = \operatorname{codim} R(T^*)$ and $\operatorname{codim} R(T) = \dim N(T^*)$.

Proof of Lemma 6.3. If $S \subseteq T$ then $N(S) \subseteq N(T)$ and $R(S) \subseteq R(T)$, so clearly $\operatorname{ind}(S) \leqq \operatorname{ind}(T)$. Equality implies that $N(S) = N(T)$ and $R(S) = R(T)$, so $S = T$.

Proof of Lemma 6.4. (a) Since $\dim N(T) < \infty$ and $\operatorname{codim} R(T) < \infty$, there are bounded projections P mapping X onto $N(T)$ and $I - Q$ mapping Y onto $R(T)$. The operator $S = (I-P)T^{-1}(I-Q)$ is closed and everywhere defined, hence bounded. Now

$$TS = I - Q \text{ and } (T-C)S = I - Q - CS.$$

Since C is T-compact and Q is finite-dimensional, Q and CS are compact. By the Riesz theory, therefore, $I - Q - CS$ is in $\Phi(Y,Y)$ and has index 0. Consequently $R(T-C) \supseteq R((T-C)S)$ is closed and has finite codimension. A purely algebraic argument shows that $\operatorname{ind}(AB) = \operatorname{ind}(A) + \operatorname{ind}(B)$. Therefore

$$\begin{aligned} \operatorname{ind}(T) &= \operatorname{ind}(TS) - \operatorname{ind}(S) \\ &= \operatorname{ind}(I-Q) - \operatorname{ind}(S) = -\operatorname{ind}(S); \\ \operatorname{ind}(T-C) &= \operatorname{ind}((T-C)S) - \operatorname{ind}(S) \\ &= \operatorname{ind}(I-Q-CS) - \operatorname{ind}(S) = -\operatorname{ind}(S). \end{aligned}$$

By Lemma 5.2, finally, $T - C$ is closed. Therefore $T - C \in \Phi(X,Y)$ and $\operatorname{ind}(T-C) = \operatorname{ind}(T)$.

(b) Let P, Q, and S be as in (a). If C is (T,ϵ)-bounded for small enough ϵ, then $\| CS \| < 1$. In this case $I - CS$ is invertible. Again Q is compact, so by (a), $(T-C)S = I - CS - Q \in \Phi(Y,Y)$, and $\operatorname{ind}((T-C)S) = 0$. The calculation of the index proceeds as in (a).

(c) C^* is closed and T^*-compact. By Lemmas 6.2 and 6.4(a), $T^* - C^* \in \Phi(Y^*, X^*)$ and $\operatorname{ind}(T^* - C^*) = \operatorname{ind}(T^*)$. Also, $D(C^*) \supseteq D(T^*)$, so $D(C^*)$ is dense, so C is closable. Therefore we also have $T - C \in \Phi(X,Y)$

and $\operatorname{ind}(T-C)=\operatorname{ind}(T)$. By Lemma 6.2, $(T-C)^*\in\Phi(Y^*,X^*)$ and $\operatorname{ind}((T-C)^*)=-\operatorname{ind}(T-C)=-\operatorname{ind}(T)=\operatorname{ind}(T^*)=\operatorname{ind}(T^*-C^*)$. Clearly $T^*-C^*\subseteq(T-C)^*$, so by Lemma 6.3 $T^*-C^*=(T-C)^*$.

COROLLARY. *Suppose that for $0\leqq\lambda\leqq 1$, $T_\lambda\colon X\to Y$ is continuous and is in $\Phi(X,Y)$. Suppose that $\lambda\to T_\lambda$ is continuous with respect to the uniform operator topology. Then* $\operatorname{ind}(T_\lambda)$ *is constant.*

Proof. For any λ, the T_λ-norm on X is equivalent to the X-norm, since T_λ is bounded. Therefore if $\lambda_n\to\lambda$, $T_{\lambda_n}\to T_\lambda$ with respect to the T_λ-norm. By Lemma 6.4(b) $\operatorname{ind}(T_\mu)=\operatorname{ind}(T_\lambda)$ for μ near λ. Therefore $\lambda\to\operatorname{ind}(T_\lambda)$ is continuous from $[0,1]$ to the integers; it must be constant. Q.E.D.

Throughout the remainder of this chapter we shall assume that G, Γ, A, B, A', and B' satisfy conditions (S1)-(S6) of Chapter 4. In addition we assume that G is bounded. Let the spaces and operators associated with (A,B) and (A',B') be defined as in Chapter 4.

THEOREM 6.1. *Let G, Γ, A, B, A', B' satisfy the conditions* (S1)-(S6) *of Chapter 4 and let G be bounded. Let the operators S and S' be defined as in Theorem 4.1. Then S and S' are Fredholm operators.*

Proof. The imbedding of Ω into $W^{2p-1,2}(G)$ is continuous and the imbedding of the latter space into $L^2(G)=E$ is compact by Proposition 1.2. Therefore the imbedding of Ω into E is compact. Also $u\to\gamma_0 u$ is continuous from Ω to $\Sigma\oplus W^{k,2}(\Gamma)$ $(k\in\Lambda_1')$ and the imbedding of $W^{k,2}(\Gamma)$ into $W^{k-2p+1}(\Gamma)$ is compact by Proposition 1.3. Therefore the imbeddding $D(S)\to F_1$ is compact. By Lemma 1, $R(S)$ is closed and $N(S)$ has finite dimension. By symmetry the same is true for $S'=S^*$. Then $\operatorname{codim} R(S)=\dim N(S')<\infty$ and $\operatorname{codim} R(S')=\dim N(S)<\infty$, so S and S' are Fredholm operators. Q.E.D.

THEOREM 6.2. *Let G, Γ, A, B, A', B' satisfy the conditions* (S1)-(S6) *of Chapter 4 and let G be bounded. Let the spaces Ω, E_1, E_2, F_1, F_2 and the operators S, γ_0, γ_1 be defined as in Chapter 4. Let $C\colon E_1\to E_2$ be such that $D(C)\subseteq\gamma_0(\Omega)$ and let $C_1\colon F_1\to F_2$ be the induced operator. Let $A(C)$ be the restriction of A to those functions u in Ω such that $\gamma_1 u'=C\gamma_0 u$. If C_1 is either S-compact and closable or else (S,ϵ)-bounded for a suitable $\epsilon>0$, then $A(C)$ is a Fredholm operator with the same index as S. If the resolvent set $r(A(C))$ is not void, then the resolvent operator is compact.*

Proof. By Lemma 4, either condition implies that $S-C_1$ is a Fredholm operator and $\operatorname{ind}(S-C_1)=\operatorname{ind}(S)$. By Theorem 5.3 $A(C)$ is closed and

densely defined. Now $D(A(C))\subseteq\Omega$ and $\Omega\to E$ is compact, so $R(A(C))$ is closed. Clearly $\dim N(A(C))=\dim N(S-C_1)$, so we need only show that $\operatorname{codim} R(A(C))=\operatorname{codim} R(S-C_1)$. But we have $\operatorname{codim} R(S-C_1)=\dim N((S-C_1)^*)=m<\infty$. Let $\{[v_1,g_1],\cdots,[v_m g_m]\}$ be a basis for $N((S-C_1)^*)$. A function u is in $R(A(C))$ if and only if $[u,0]\in R(S-C_1)$, and $[u,0]\in R(S-C_1)$ if and only if $\langle[u,0],[v_j,g_j]\rangle=0$ or $\langle u,v_j\rangle=0$, $1\leqq j\leqq m$. Therefore the v_j span $N(A(C)^*)$. It remains to be shown that the v_j are linearly independent. If not, there would be a $[0,g]\in N((S-C_1)^*)$ with $g\neq 0$. Then for all $u\in\Omega$,

$$0=\langle(S-C_1)[u,\gamma_0 u],[0,g]\rangle=\langle\gamma_1 u-C\gamma_0 u,g\rangle.$$

But by Lemma 3.1, $\{\gamma_1 u\mid u\in\Omega,\gamma_0 u=0\}$ is dense in E_2. Therefore $g=0$, the v_j are independent, and $\operatorname{codim} R(A(C))=\operatorname{codim} R(S-C_1)$.

Suppose $\lambda\in r(A(C))$. Clearly λI is $A(C)$-compact, so by Lemma 5.1 the $(A(C)-\lambda I)$-norm on $D(A(C))$ is equivalent to the $A(C)$ norm. The resolvent operator is therefore continuous as an operator from E to $D(A(C))$ with the $A(C)$ norm, hence compact as an operator in E. This completes the proof of Theorem 1.

COROLLARY. *Let $C: E_1\to E_2$ and $C^*: E_1'\to E_2'$ satisfy the conditions of Theorem 2 with respect to S and S' respectively. Then $A(C)$ and $A'(C^*)$ are adjoint.*

Proof. By Lemma 6.4, $(S-C_1)^*=S'-C_1^*$. Therefore Theorem 5.4 applies.

In some cases we can determine the index of S a priori. We need one further assumption:

(S7) The coefficients of the principal parts of A, A', B_k and B_k' are in $C^{2p}(G\cup\Gamma)$ and $C^{2p}(K)$, respectively, for $k=0,1,\cdots,2p-1$.

THEOREM 6.3. *Suppose* (S7) *is satisfied.*

(a) *Suppose $\Lambda_1=\Lambda_1'$ and the principal part of γ_1' is the complex conjugate of the principal part of γ_1. Then* $\operatorname{ind}(S)=\operatorname{ind}(S')=0$.

(b) *Suppose γ_1 and γ_1' are both of Dirichlet type. Then* $\operatorname{ind}(S)=\operatorname{ind}(S')=0$.

Proof. (a) Let A_0 be the principal part of A and δ_1 the principal part of γ_1. Let $N_0,N_1,\cdots,N_{2p-1}$ be the normal derivatives: $N_k=i^{-k}\dfrac{\partial^k}{\partial n^k}$. Put

$$\delta_0=(i^k N_k;k\in\Lambda_0).$$

Then (A_0,δ_0,δ_1) and its conjugate system satisfy the conditions (S1)-(S6),

so the associated operator S_0 is a Fredholm operator. By the corollary of Theorem 4.1, $\Omega(S_0)=\Omega(S)$. Let $T:\Omega\to F_2$ and $T_0:\Omega\to F_2$ be defined by $Tu=[Au,\gamma_1 u]$ and $T_0 u=[A_0 u,\delta_1 u]$. T and T_0 differ by operators of lower order, so $T-T_0$ is compact as a mapping from Ω to F_2. T and T_0 are clearly Fredholm operators, so $\operatorname{ind}(T)=\operatorname{ind}(T_0)$. Now $R(T)=R(S)$ and $\dim N(T)=\dim N(S)$, so $\operatorname{ind}(T)=\operatorname{ind}(S)$. Likewise $\operatorname{ind}(T_0)=\operatorname{ind}(S_0)$, so $\operatorname{ind}(S)=\operatorname{ind}(S_0)$.

Similarly, let A_0' and δ_1' be the principal parts of A' and δ_1'. Let S_0' correspond to $(A_0',\delta_0',\delta_1')$. By the argument above, $\operatorname{ind}(S')=\operatorname{ind}(S_0')$. Since $\operatorname{ind}(S)=-\operatorname{ind}(S')$, we can complete the proof by showing that $\operatorname{ind}(S_0)=\operatorname{ind}(S_0')$. Let $J:F_1\to F_1'$ be defined by $J[u,f]=[\bar{u},\bar{f}]$. Then J is an isometry and a conjugate linear mapping of F_1 onto itself. Since $\overline{\delta_0 u}=\delta_0\bar{u}$, by the corollary to Theorem 4.1, J maps $D(S_0)$ onto $D(S_0')$. The operators A_0',δ_0',δ_1' are the complex conjugates of A,δ_0,δ_1, so $S_0'J=JS_0$. Therefore $\operatorname{ind}(S_0')=\operatorname{ind}(S_0'J)=\operatorname{ind}(JS_0)=\operatorname{ind}(S_0)$. This proves (a).

(b) Suppose γ_1 and γ_1' are of Dirichlet type. Then necessarily $\Lambda_1=\Lambda_1'=(0,1,\cdots,p-1)$ or $\Lambda_1=\Lambda_1'=(p,p+1,\cdots,2p-1)$. Let δ_0 be defined as in (a) and let

$$\delta_1=(i^k N_k;k\in\Lambda_1).$$

Since γ_1 is normal there is a vector function $b(x)=(b_k(x);k\in\Lambda_1)$ such that $b_k(x)\in C^{2p}(\Gamma)$ and $b_k(x)\neq 0$, $x\in\Gamma$, $k\in\Lambda_1$, while $b(x)\gamma_1-\delta_1$ involves only tangential derivatives. (Here $b(x)\gamma_1=(b_k(x)B_k;k\in\Lambda_1)$. Let $\gamma_{1,0}$ be the principal part of γ_1. For $0\leqq\lambda\leqq 1$, put

$$\delta_{1,\lambda}=(1-\lambda)b(x)\gamma_{1,0}+\lambda\delta_1.$$

For each λ, $\delta_{1,\lambda}$ is of Dirichlet type. Let A_0 be the principal part of A. Then $(A_0,\delta_0,\delta_{1,\lambda})$ and its conjugate satisfy (S1)-(S6) for every λ, so the associated operator S_λ is a Fredholm operator. $D(S_\lambda)\equiv D(S_0)$. Considered as an operator from $D(S_0)$ to F_2, each S_λ is closed and everywhere defined, hence continuous. Because of the form of $\delta_{1,\lambda}$ $\lambda\to S_\lambda$ is continuous with respect to the appropriate uniform operator topology, so by the corollary of Lemma 4, $\operatorname{ind}(S_0)=\operatorname{ind}(S_1)$. S_0 corresponds to $(A_0,\delta_0,b(x)\gamma_{1,0})$, so $\operatorname{ind}(S_0)$ is the same as the index of the operator $S_{0,1}$ corresponding to $(A_0,\delta_0,b(x)\gamma_1)$. But $\dim N(S_{0,1})=\dim N(S)$ and $[u,f]\to[u,b(x)f]$ is a topological isomorphism taking $R(S)$ onto $R(S_{0,1})$. Therefore

$$\operatorname{ind}(S)=\operatorname{ind}(S_{0,1})=\operatorname{ind}(S_0)=\operatorname{ind}(S_1).$$

S_1 corresponds to (A_0,δ_0,δ_1). Similarly $\operatorname{ind}(S')=\operatorname{ind}(S_1')$, where S_1' cor-

responds to $(A_0', \delta_0, \delta_1)$ and A_0' is the principal part of A'. The operators A_0', δ_0, δ_1 are the complex conjugates of A_0, δ_0, δ_1, so as in (a), $\operatorname{ind}(S_1) = \operatorname{ind}(S_1')$. Therefore $\operatorname{ind}(S) = \operatorname{ind}(S') = 0$. Q. E. D.

Theorems 2 and 3(a) give a generalization to the realizations $A(C)$ of a result of Kaniel and Schechter [20]. Note that by either 3(a) or 3(b) the Dirichlet realization has index 0.

Chapter 7. Second Order Operators.

Let G, Γ, A, B, A', and B' satisfy the conditions (S1)-(S5) of Chapter 4, with $p = 1$ and $\Lambda_1 = \Lambda_1' = \{1\}$. Then the spaces and operators are

$$E_1 = E_2 = E_1' = E_2' = L^2(\Gamma);$$
$$F_1 = F_2 = F_1' = F_2' = L^2(G) \oplus L^2(\Gamma);$$
$$\gamma_0 = B_0, \quad \gamma_1 = B_1, \quad \gamma_0' = B_0', \quad \gamma_1' = B_1'.$$

We can summarize some of the previous results for this case:

THEOREM 7.1. *Let C and C' be operators in $L^2(\Gamma)$ whose domains include $W^{1,2}(\Gamma)$. Let C_1 and C_1' be the induced operators in $L^2(G) \oplus L^2(\Gamma)$ and let $A(C)$ and $A'(C')$ be the corresponding realizations of A and A'.*

(a) *If $R(C) \subseteq W^{1,2}(\Gamma)$, then $D(A(C)) \subseteq W^{2,2}(G)$.*

(b) *If C is bounded as an operator from $W^{s,2}(\Gamma)$ to $L^2(\Gamma)$ for some $s < 1$, then $A(C)$ is closed.*

(c) *If C is compact from $W^{1,2}(\Gamma)$ to $L^2(\Gamma)$ and C_1 is closable, then $A(C)$ is closed.*

(d) *If C and C' each satisfy the hypotheses either of* (a) *and* (b) *or of* (c), *and if $(S - C_1)^* = S' - C_1'$, then $A(C)$ and $A'(C')$ are adjoints.*

(e) *If G is bounded and C satisfies the hypotheses of* (c), *then $A(C)$ is a Fredholm operator with* $\operatorname{ind}(A(C)) = \operatorname{ind}(S)$. *If the resolvent set $r(A(C))$ is not empty, the resolvent operator is compact.*

(f) *If G is bounded, C and C' satisfy the hypotheses of* (c), *and $C' \subseteq C^*$, then $A(C)$ and $A'(C')$ are adjoints.*

Proof. Parts (a) and (b) follow from the methods of Theorems 5.4 and 5.6, as does the corresponding portion of part (d). Part (c) and the rest of (d) follow from 5.5. Parts (e) and (f) are just Theorem 6.2 and its corollary.

COROLLARY. *If Γ is bounded and C is any bounded operator in $L^2(\Gamma)$, then the realizations $A(C)$ and $A'(C^*)$ are adjoints.*

Proof. If Γ is bounded the injection $W^{1,2}(\Gamma) \to L^2(\Gamma)$ is compact. Therefore C and C^* are compact as mappings from $W^{1,2}(\Gamma)$ to $L^2(\Gamma)$. Since C_1 is bounded, $(S-C_1^*) = S^* - C_1^* = S' - C_1^*$. Thus Theorem 1(d) applies.

Throughout the remainder of this chapter we shall assume that G and Γ satisfy the conditions of Chapter 4, for $p=1$. Let $A = \sum_{|\alpha|\leq 2} a_\alpha D^\alpha$, where each a_α is continuous on $G \cup \Gamma$ and in $L^\infty(G)$. Assume a_α has bounded continuous derivatives of order ≤ 2 on $G \cup \Gamma$ for $|\alpha| = 2$, and a_α has bounded continuous first derivatives on G for $|\alpha| = 1$. Then A can be written as $A_0 + L$, where

$$A_0 = \sum D_j(a_{jk}D_k), \qquad L = \sum b_j D_j + b_0.$$

We may assume $a_{jk} = a_{kj}$, the a_{kj} have bounded continuous derivatives of order ≤ 2, the b_j have bounded continuous first derivatives, and $b_0 \in L^\infty(G)$. The formal adjoint of A is $A_0' + L'$, where

$$A_0' = \sum D_j(\bar{a}_{jk}D_k), \qquad L' = \sum D_j \bar{b}_j + \bar{b}_0.$$

Associated with A is the continuous bilinear form $D(u,v)$ defined on $W^{1,2}(G)$ by

$$D(u,v) = \sum (a_{jk}D_k u, D_j v)_G.$$

Associated with A' is the form $D'(v,u)$, where

$$D'(v,u) = \sum (\bar{a}_{jk}D_k v, D_j u)_G = \overline{D(u,v)}.$$

Suppose $u, v \in C_0^\infty(G \cup \Gamma)$. Then

$$(A_0 u, v)_G - D(u,v) = \sum (i n_j a_{jk} D_k u, v)_\Gamma,$$

or

$$(A_0 u, v)_G = D(u,v) - (\delta_1 u, \delta_0 v)_\Gamma, \tag{1}$$

where

$$\delta_0 v = v\,|\,\Gamma, \qquad \delta_1 u = -i \sum n_j a_{jk}(D_k u\,|\,\Gamma). \tag{2}$$

Similarly

$$(A_0' v, u)_G = D'(v,u) - (\delta_1' v, \delta_0 u)_\Gamma, \tag{3}$$

where

$$\delta_1' v = -i \sum n_j \bar{a}_{jk}(D_k u\,|\,\Gamma). \tag{4}$$

By continuity (1) holds for all $u \in \Omega$ and all $v \in W^{1,2}(G)$, while (3) holds for all $v \in \Omega'$ and $u \in W^{1,2}(G)$. If $u \in \Omega$ and $v \in \Omega'$, we can subtract the complex conjugate of (3) from (1) to get

$$(5) \qquad (A_0u, v)_G - (u, A_0'v)_G = (\delta_0 u, \delta_1' v)_\Gamma - (\delta_1 u, \delta_0 v)_\Gamma.$$

If $u \in \Omega$ and $v \in \Omega'$,

$$(6) \qquad \begin{aligned} &(Au, v)_G - (u, A'v)_G \\ &\qquad = (A_0u, v)_G - (u, A_0'v)_G + (Lu, v)_G - (u, L'v)_G \\ &\qquad = (\delta_0 u, \delta_1' v)_\Gamma - (\delta_1 u, \delta_0 v)_\Gamma + (c\delta_0 u, \delta_0 v)_\Gamma, \end{aligned}$$

where $c(s) = i\sum n_j(x) b_j(x)$. Let $b(x)$ and $b'(x)$ be two functions in $C'(G \cup \Gamma)$ with bounded first derivatives, such that $b - \bar{b}' = c$; e. g. take $b = Re(c)$ and $b' = iIm(c)$, where $Re(c)$ and $Im(c)$ denote the real and imaginary parts. Let

$$(7) \qquad \gamma_0 = \gamma_0' = \delta_0, \qquad \gamma_1 = \delta_1 - b\delta_0, \qquad \gamma' = \delta_1 - b'\delta_0.$$

Then the systems (A, γ_0, γ_1) and $(A', \gamma_0', \gamma_1')$ are conjugate. In what follows we shall assume that A is uniformly elliptic and regularly elliptic on G.

LEMMA 7.1. *The systems (A, γ_1) and (A', γ_1') are uniformly regular.*

Proof. Let x be a point of Γ, n the normal at x, and t a tangent vector at x. Let λ_1 and λ_2 be the roots of $a(x, t + \lambda n)$ having positive and negative imaginary parts, respectively. Let C be a circle in the upper half plane with center λ_1. Let $A(\xi, \eta) = \sum a_{jk}(x)\xi_j\eta_k$. This form is symmetric, so

$$\begin{aligned} a(x, t + \lambda n) &= A(t + \lambda n, t + \lambda n) \\ &= \lambda^2 A(n, n) + 2\lambda A(n, t) + A(t, t). \end{aligned}$$

Therefore

$$(9) \qquad \lambda_1 = -A(n, t)/A(n, n) + \tfrac{1}{2}(\lambda_1 - \lambda_2).$$

Equation (2) gives $b(x, \xi) = -iA(n, \xi)$. Consequently

$$(10) \qquad \begin{aligned} \int_C \frac{b(x, t + \lambda n)}{a(x, t + \lambda n)}\, d\lambda &= \frac{1}{A(n, n)} \int_C \frac{b(x, t + \lambda n)}{(\lambda - \lambda_1)(\lambda - \lambda_2)}\, d\lambda \\ &= \frac{2\pi i}{A(n, n)(\lambda_1 - \lambda_2)}\, b(x, t + \lambda_1 n) \\ &= \frac{2\pi i}{A(n, n)(\lambda_1 - \lambda_2)}\, (-iA(n, t + \lambda_1 n)) \\ &= \frac{2\pi i}{A(n, n)(\lambda_1 - \lambda_2)}\, (-i/2[A(n, t) + A(n, n)\lambda_1]). \end{aligned}$$

Combining (9) and (10) we get

$$\int_C \frac{b(x,t+\lambda n)}{a(x,t+\lambda n)}\, d\lambda = \pi. \tag{11}$$

Since this is true for all x and t, (A,γ_1) is uniformly regular. By symmetry so is (A',γ_1'). Q.E.D.

COROLLARY. *(A,γ_0,γ_1) and (A,γ_0',γ_1') satisfy the conditions* (S1)-(S6). *If G is bounded, the associated operators S and S' have index* 0.

Proof. This follows from the lemma and Theorem 6.3.

In case A is uniformly strongly elliptic or essentially real the techniques of Bade and Freeman [3] can be used to improve Theorem 1 for (A,γ_0,γ_1) and to obtain estimates on the location of the spectra af realizations. The operator A is said to be ***strongly elliptic*** if the characteristic polynomial $a(x,\xi)$ has real part $\operatorname{Re} a(x,\xi) > 0$ for all $\xi \neq 0$ in R^n. A is said to be ***uniformly strongly elliptic*** if there is a constant $c_0 > 0$ such that $\operatorname{Re} a(x,\xi) \geqq c_0 |\xi|^2$, all $\xi \in R^n$, $x \in G$.

LEMMA 7.2. *A is uniformly strongly elliptic, then for all $u \in W^{1,2}(G)$, $\operatorname{Re} D(u,u) \geqq c_0(\|u\|_{1,G})^2 - c_0(\|u\|_{0,G})^2$.*

Proof. Since $a_{jk} = a_{kj}$,

$$\begin{aligned}\operatorname{Re} D(u,u) &= \tfrac{1}{2}[D(u,u) + D'(u,u)]\\ &= \tfrac{1}{2}\sum ((a_{jk}+\bar{a}_{jk})D_k u, D_j v)_G\\ &= \int_G \operatorname{Re}[a(x,(D_1u,\cdots,D_nu))]\,dx\\ &\geqq c_0 \int_G \sum |D_j u|^2\,dx = c_0(\|u\|_{1,G})^2 - c_0(\|u\|_{0,G})^2.\end{aligned}$$

LEMMA 7.3. *There is a constant $M_0 > 0$ such that for all u in $W^{1,2}(G)$,*

$$\|\gamma_0 u\|^2_{0,\Gamma} \leqq M_0(\|u\|_{1,G}\,\|u\|_{0,G}).$$

Proof. As usual, let $n(x) = (n_1(x),\cdots,n_n(x))$ be the unit inner normal to Γ at x. For some $\epsilon_0 > 0$ the mappings $\pi_\epsilon(x) = x + \epsilon n(x)$, $0 \leqq \epsilon \leqq \epsilon_0$, are homeomorphisms of Γ onto Γ_ϵ which are uniformly of class 2. Define functions $\phi_1(x),\cdots,\phi_n(x)$ on $G \cup \Gamma$ by putting

$$\phi_j(x) = (1-\epsilon/\epsilon_0)^2 n_j(\pi_\epsilon^{-1}(x)),\ x \in \Gamma_\epsilon,\ 0 \leqq \epsilon \leqq \epsilon_0.$$
$$\phi_j(x) = 0 \text{ otherwise.}$$

Then the $\phi_j(x)$ are in $C'(G\cap\Gamma)\cap L^\infty(G)$ and have bounded first derivatives, while $\gamma_0\phi_j=n_j$. Suppose $u\in C_c^\infty(G\cup\Gamma)$. Then

$$\begin{aligned}(\|\gamma_0 u\|_{0,\Gamma})^2&=\int_\Gamma\ (\sum_{j=1}^n n_j^2(x))\,|u(x)|^2\,dx\\&=-\sum_{j=1}^n\int_G\frac{\partial}{\partial x_j}(\phi_j(x)\,|u(x)|^2)\,dx\\&=\sum_{j=1}^n i[(D_j(\phi_j u),u)_G-(\phi_j u,D_j u)_G]\\&\leqq M_0\,\|u\|_{1,G}\,\|u\|_{0,G}.\end{aligned}$$

By continuity, the inequality holds for all $u\in W^{1,2}(G)$.

THEOREM 7.2. *Suppose A is uniformly strongly elliptic. Let C and C' be operators in $L^2(G)$ such that C and C' are each bounded with norm $\leqq N$ as operators from $W^{\frac12,2}(\Gamma)$ to $L^2(\Gamma)$. Let C_1 and C_1' be the induced operators in $L^2(G)\oplus L^2(\Gamma)$, and suppose $(S-C_1)^*=S'-C_1'$.*

Then $A(C)$ and $A'(C')$ are adjoints. Furthermore the spectrum of $A(C)$ and the spectrum of $A'(C')$ are each contained in a triangular region

$$\operatorname{Re}(\lambda)\geqq-c(1+N)^4+d\,|\operatorname{Im}(\lambda)|,$$

where c and d are positive constants independent of N.

Proof. Theorem 1 does not apply, since we do not assume (a). We wish to apply Theorem 5.1, and must show that $S-C_1$ and $S'-C_1'$ satisfy conditions (T1)-(T5) and (T1)'-(T5)'. Certainly (T1) and (T1)' are satisfied. The assumptions on C and C' together with Proposition 1.4, the corollary to Proposition 1.1, and Lemma 6.1 imply (T2), (T3), (T2)', and (T3)'. Condition (T5) is assumed, and (T5)' follows from (T2) and (T5). This leaves (T4) and (T4)'. By symmetry we need only show T4, and this will follow if $(\gamma_1-C\gamma_0)(\Omega)$ is closed.

For complex λ, let P_λ be the bounded operator in $L^2(G)\oplus L^2(\Gamma)$ given by $P_\lambda[u,f]=[\lambda u,0]$. If $u\in\Omega$ then

$$\begin{aligned}&\langle(S-P_\mu-C_1)[u,\gamma_0 u],[u,\gamma_0 u]\rangle\\&\quad=(Au,u)+(\gamma_1 u,\gamma_0 u)-\mu\,\|u\|_0^2-(C\gamma_0 u,\gamma_0 u)\\&\quad=(A_0u,u)+(\delta_1 u,\delta_0 u)-\mu\,\|u\|_0^2+R(u)\\&\quad=D(u,u)-\mu\,\|u\|_0^2+R(u),\end{aligned}\tag{12}$$

where

$$R(u)=(Lu,u)-(b\gamma_0 u,\gamma_0 u)-(C\gamma_0 u,\gamma_0 u).\tag{13}$$

(For convenience we drop the subscripts G and Γ on the inner products.) By Proposition 1.4 and Lemma 3,

$$(14) \qquad |R(u)| \leqq M_1 \| u \|_1 \| u \|_0 + M_2 N \| u \|_1 \| \gamma_0 u \|_0.$$

Note that if $a \geqq 0$, $b \geqq 0$, and $t > 0$ then

$$ab = (at)(b/t) \leqq t^2 a^2 + t^{-2} b^2.$$

Applying this to (14) twice we get

$$(15) \qquad \begin{aligned} &|R(u)| \\ &\leqq (\epsilon \| u \|_1^2 + M_1^2 \epsilon^{-1} \| u \|_0^2) + (\epsilon \| u \|_1^2 + \epsilon^{-1} M_2^2 N^2 \| \gamma_0 u \|_0^2) \\ &\leqq 3\epsilon \| u \|_1^2 + (M_4 \epsilon^{-1} + M_3 \epsilon^{-3} N^4) \| u \|_0^2. \end{aligned}$$

Therefore if $u \in \Omega$, μ is real, and $\epsilon > 0$,

$$(16) \qquad \begin{aligned} &\operatorname{Re}\langle (S - P_\mu - C_1)[u, \gamma_0 u], [u, \gamma_0 u] \rangle \\ &= \operatorname{Re} D(u, u) - \mu \| u \|_0^2 - \operatorname{Re}(R(u)) \\ &\geqq (c_0 - 3\epsilon) \| u \|_1^2 + (-\mu - c_0 - M_3 \epsilon^{-1} - M_4 N^4 \epsilon^{-3}) \| u \|_0^2. \end{aligned}$$

Let $\epsilon = c_0/6$, $c_1 = M_3/\epsilon + c_0$, $d_1 = M_4 \epsilon^{-3}$. Then if $\mu < -c_1 - d_1 N^4$, (16) gives

$$(17) \qquad |\langle (S - P_\mu - C_1)[u, \gamma_0 u], [u, \gamma_0 u] \rangle| \geqq c_0/2 \| u \|_1^2.$$

For some $e_0 > 0$, $\| u \|_1^2 \geqq e_0 \| [u, \gamma_0 u] \|^2$. Then (17) implies

$$(18) \qquad \| (S - P_\mu - C_1)[u, \gamma_0 u] \| \geqq e_0 c_0/2 \| [u, \gamma_0 u] \|^2.$$

Therefore $S - P_\mu - C_1$ has trivial null space and closed range. By symmetry, for small enough μ the same is true of $S' - P_\mu - C_1' = (S - C_1 - P_\mu)^*$. Therefore $R(S - P_\mu - C_1) = L^2(G) \oplus L^2(\Gamma)$, so $(\gamma_1 - C\gamma_0)(\Omega) = L^2(\Gamma)$ and (T4) is satisfied. This completes the proof that $A(C)$ and $A'(C')$ are adjoints.

Finally we wish to determine the location of the spectra of $A(C)$ and $A'(C')$. By symmetry and the above argument it is sufficient to show that for λ outside a set of the above form there is an $e_0 = e_0(\lambda) > 0$ such that for $u \in D(A(C))$,

$$(19) \qquad |(Au - \lambda u, u)| \geqq e_0 \| u \|_0^2.$$

Note that if $a \geqq 0$, $b \geqq 0$, and $0 \leqq t \leqq 1$ then

$$\begin{aligned} (a^2 + b^2)^{\frac{1}{2}} &= (t + (1 - t))^{\frac{1}{2}} (a^2 + b^2)^{\frac{1}{2}} \\ &\geqq (t^2 + (1 - t)^2)^{\frac{1}{2}} (a^2 + b^2)^{\frac{1}{2}} \geqq ta + (1 - t) b. \end{aligned}$$

Suppose $u \in D(A(C))$. Then

$$(Au - \lambda u, u) = D(u, u) - \lambda \| u \|_0^2 + R(u), \tag{20}$$

where $R(u)$ is given by (13). Therefore if $0 \leqq t \leqq 1$ and $\lambda = \mu + i\nu$, with μ and ν real.

$$\begin{aligned} |(Au - \lambda u, u)| \geqq t \operatorname{Re} D(u, u) - t\mu \| u \|_0^2 - | R(u) | \\ + (1 - t) | \nu | \| u \|_0^2 - (1 - t) | \operatorname{Im} D(u, u) |. \end{aligned} \tag{21}$$

Using (15) and the fact that $D(u, v)$ is a bounded form on $W^{1,2}(G)$ we get an inquality of the form

$$\begin{aligned} |(Au - \lambda u, u)| \geqq [tc_0 - \epsilon - (1 - t) M_5] \| u \|_1^2 \\ + [| \nu | - t\mu - (M_6 + M_7 N^4) \epsilon^{-3}] \| u \|_0^2. \end{aligned} \tag{22}$$

Take $t = 2M_5/(2M_5 + c_0)$ and $\epsilon = c_0 t/2$. Then the desired inequality holds if $| \nu | > (t\mu + (M_6 + M_7 N^4)) \epsilon^{-3} (1 - t)^{-1}$, or

$$\mu < - d_2 (1 + N)^4 + c_2 | \nu |, \tag{23}$$

for suitable c_2 and d_2. This completes the proof of Theorem 2.

The operator A is said to be *essentially real* if a_{jk} is real, $1 \leqq j, k \leqq n$. If A is essentially real, then for any $x_0 \in G$, $a(x_0, \xi)$ has real coefficients. Since $a(x_0, \xi)$ has no non-zero real roots this implies that either $a(x_0, \xi) > 0$, all $\xi \neq 0$ in R^n, or $a(x_0, \xi) < 0$, all $\xi \neq 0$ in R^n. The inequality must hold in the same direction for all x in one component of G. Therefore if G is connected and A is essentially real, either A or $-A$ is strongly elliptic.

THEOREM 7.3. *Suppose A is uniformly strongly elliptic and essentially real. Suppose $0 \leqq \eta \leqq \frac{1}{2}$, and let C and C' be operators in $L^2(G)$ which are each bounded with norm $\leqq N$ as operators from $W^{\frac{1}{2}-\eta,2}(\Gamma)$ to $L^2(\Gamma)$. Let C_1 and C_1' be the induced operators in $L^2(G) \oplus L^2(\Gamma)$ and suppose $(S - C_1)^* = S' - C_1'$.*

Then $A(C)$ and $A'(C')$ are adjoints. The spectrum of $A(C)$ and the spectrum of $A'(C')$ are each contained in a region of the form

$$\operatorname{Re}(\lambda) \geqq - c(1 + N)^4 + d | \operatorname{Im}(\lambda) |^s (1 + N)^{-(1+2\eta)s},$$

where c and d are positive constants independent of N, and $s = 4/(3 - 2\eta)$.

Proof. The first part of the conclusion follows from Theorem 2. As in the proof of Theorem 2 it suffices by symmetry to show that for λ outside a region of the above form, the inequality (19) holds for some $e_0 = e_0(\lambda) > 0$.

The procedure is similar to that in Theorem 2. Since A is essentially real, $D(u,u)$ is real. Equation (21) becomes

$$(24) \qquad |(Au-\lambda u,u)| \geqq tD(u,u) - t\mu \| u \|_0^2 + (1-t)|\nu| - |R(u)|,$$

where $\lambda = u + i\nu$ and $0 \leqq t \leqq 1$.

Suppose $0 < \eta \leqq \frac{1}{2}$. Under the above hypotheses by an argument similar to that in Proposition 1.4 we get

$$(25) \qquad |R(u)| \leqq M_1 N \| u \|_{1-\eta} \| \delta_0 u \|_0 + M_2 \| u \|_1 \| u \|_0.$$

The corollary to Proposition 1.1 shows that for $0 < \epsilon \leqq \epsilon(\eta) < 1$, we have

$$(\| u \|_{1-\eta})^2 \leqq \epsilon \| u \|_1^2 + \epsilon^{-\alpha} \| u \|_0^2,$$

where $\alpha = (1-\eta)/\eta$. For $\beta > 0$ and $0 < \epsilon \leqq \epsilon(\eta)$,

$$\begin{aligned} M_1 N \| u \|_{1-\eta} \| \delta_0 u \|_0 &\leqq \epsilon^{\beta} \| u \|_{1-\eta}^2 + \epsilon^{-\beta} M_1^2 N^2 \| \delta_0 u \|_0^2 \\ &\leqq \epsilon^{\beta} (\| u \|_{1-\eta})^2 + \epsilon^{-\beta} M_2 N^2 \| u \|_1 \| u \|_0 \\ &\leqq (\epsilon \| u \|_1^2 + \epsilon^{\beta-\alpha(1-\beta)} \| u \|_0^2) \\ &\quad + (\epsilon \| u \|_1^2 + \epsilon^{-\beta-(\beta+1)} M_2^2 N^4 \| u \|_0^2). \end{aligned}$$

Set $\beta = (\alpha-1)/(\alpha+3)$, so that $\beta - \alpha(1-\beta) = 2\beta - 1 = r$, where $r = (3-2\eta)/(1+2\eta)$. Then the last inequality is

$$(26) \qquad M_1 N \| u \|_{1-\eta} \| \delta_0 u \|_0 \leqq 2\epsilon \| u \|_1^2 + (M_3 N^4 + 1)\epsilon^{-r} \| u \|_0^2.$$

Note that $r = r(\eta) \to 3$ as $\eta \to 0$. Therefore if we take $\epsilon(0) = 1$ we can combine (25), (26), and (15) to get

$$(27) \qquad |R(u)| \leqq \epsilon \| u \|_1^2 + K_0 \epsilon^{-r} \| u \|_0^2.$$

Here $0 \leqq \eta \leqq \frac{1}{2}$, $r = (3-2\eta)/(1+2\eta)$, $0 < \epsilon \leqq \epsilon(\eta)$, and $K_0 = M_4 N^4 + M_5$, where M_4 and M_5 depend only on η.

We now combine (24) and (27) to get

$$(28) \qquad |(Au-\lambda u,u)| \geqq (c_0 t - \epsilon) \| u \|_1^2 + [(1-t)|\nu| - t\mu - K_1 \epsilon^{-r}] \| u \|_0^2,$$

where $K_1 = K_0 + c_0$, $u \in D(A(C))$, and $0 \leqq t \leqq 1$, $0 < \epsilon \leqq \epsilon(\eta)$. Suppose $0 \leqq t \leqq t(\eta)$, where $t(\eta) = \min(1, \epsilon(\eta)/c_0)$. Then (28) implies

$$(29) \qquad |(Au-\lambda u,u)| \geqq [-t\mu + (1-t)|\nu| - K_1 t^{-r}] \| u \|_0^2.$$

Thus the desired inequality (19) holds for any $0 < t \leqq t(\eta)$ and any $\lambda = \mu + i\nu$ with

$$(30) \qquad \mu < |\nu|(1-t)t^{-1} - Kt^{-(r+1)}.$$

For the best possible result we wish to maximize

$$g(t) = |\nu|(1-t)t^{-1} - Kt^{-(r+1)},$$

for $0 < t \leqq t(\eta)$. The maximum of $g(t)$ for $0 \leqq t \leqq 1$ is attained at $t(\nu) = ((r+1)K/|\nu|)^{1/r}$, and $g(t(\nu)) = O(|\nu|^{1+1/r}K^{-1/r})$ as $|\nu| \to \infty$. Therefore if $t(\nu) \geqq t(\eta)$ we can let $t = t(\eta)$, while for $|\nu|$ large we can take $t = t(\nu)$ in (30). Since $1 + 1/r = 4/(3 - 2\eta)$ and $-1/r = (1 + 2\eta)/(2\eta - 3)$, the inequality (19) does indeed hold for λ outside a region of the desired form. This completes the proof of Theorem 4.

In particular, suppose C is a bounded operator in $L^2(G)$. Then we may take $\eta = \frac{1}{2}$ in Theorem 4, and $s = 2$. Furthermore we necessarily have $(S - C_1)^* = S' - C_1^*$. Therefore we have

COROLLARY. *Let A be uniformly strongly elliptic and essentially real, and let C be any bounded operator in $L^2(\Gamma)$. Then $A(C)$ and $A'(C^*)$ are adjoints. The spectrum of $A(C)$ and the spectrum of $A'(C^*)$ are each contained in a parabolic region opening to the right. Furthermore, if G is bounded then $A(C)$ and $A'(C^*)$ are Fredholm operators with compact resolvents.*

This last result was obtained for formally self-adjoint operators by Freeman [17]; however in the case of an unbounded boundary Γ he restricts C to be multiplication by a bounded function.

YALE UNIVERSITY.

REFERENCES.

[1] S. Agmon, A. Douglis, and L. Nirenberg, "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, I," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 623-727.

[2] N. Aronszajn and A. N. Milgram, "Differential operators on Riemannian manifolds," *Rendiconti del Circolo Matematico di Palermo* (2), vol. 11 (1953), pp. 266-325.

[3] W. G. Bade and R. S. Freeman, "Closed extensions of the Laplace operator determined by a general class of boundary conditions," *Pacific Journal of Mathematics*, vol. 12 (1962), pp. 395-410.

[4] R. Beals, *Non-local boundary value problems for elliptic partial differential operators*, Dissertation, Yale, 1964.

[5] ———, "A note on the adjoint of a perturbed operator," *Bulletin of the American Mathematical Society*, vol. 70 (1964), pp. 314-315.

[6] ———, "Non-local elliptic boundary value problems," *Bulletin of the American Mathematical Society*, vol. 70 (1964), pp. 693-696.

[7] F. E. Browder, "Estimates and existence theorems for elliptic boundary value

problems," *Proceedings of the National Academy of Sciences of the U. S. A.*, vol. 45 (1959), pp. 365-372.

[8] ———, "A-priori estimates for solutions of elliptic boundary value problems, I, II, III," *Koninklijke Nederlandse Akademie van Wetenschappen, Indagationes Mathematicae*, vol. 22 (1960), pp. 145-159, 160-169; vol. 23 (1961), pp. 404-410.

[9] ———, "On the spectral theory of elliptic differential operators, I," *Mathematische Annalen*, vol. 142 (1961), pp. 22-130.

[10] ———, "Functional analysis and partial differential equations, II," *Mathematische Annalen*, vol. 145 (1962), pp. 81-226.

[11] A. P. Calderon and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), pp. 85-139.

[12] ———, "On singular integrals," *American Journal of Mathematics*, vol. 78 (1956), pp. 289-309.

[13] J. W. Calkin, "Abstract symmetric boundary conditions," *Transactions of the American Mathematical Society*, vol. 45 (1939), pp. 369-442.

[14] ———, "General self-adjoint boundary conditions for certain partial differential operators," *Proceedings of the National Academy of Sciences of the U. S. A.*, vol. 25 (1939), pp. 201-206.

[15] N. Dunford and J. T. Schwartz, *Linear Operators*, Part I, New York, 1958.

[16] R. S. Freeman, "Closed extensions of the Laplace operator determined by a general class of boundary conditions for unbounded regions," *to appear.*

[17] ———, "On closed extensions of second order formally self-adjoint uniformly elliptic differential operators," *to appear.*

[18] I. C. Gohberg and M. G. Krein, "The basic propositions on defect numbers, root numbers, and indices of linear operators," *Uspehi Matematiceskih Nauk* (*N. S.*), vol. 12 (1957), pp. 43-118, translated in *American Mathematical Society Translations*, Series 2, vol. 13 (1960), pp. 185-264.

[19] L. Hormander, *Linear Partial Differential Operators*, New York, 1963.

[20] S. Kaniel and M. Schechter, "Spectral theory for Fredholm operators," *Communications on Pure and Applied Mathematics*, vol. 16 (1963), pp. 423-448.

[21] J. L. Lions, *Lectures on Partial Differential Equations*, Tata Institute lecture notes, Bombay, 1957.

[22] J. Peetre, "Théorèmes de régularité pour quelques classes des opérateurs différentiels," *Meddelanden fran Lunds Universitets Matematiska Seminarium*, Band 16 (1959).

[23] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, New York, 1955.

[24] M. Schechter, "Integral inequalities for partial differential operators and functions satisfying general boundary conditions," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 37-66.

[25] ———, "General boundary value problems for elliptic partial differential equations," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 457-486.

[26] L. Schwartz, *Théorie des Distributions*, Paris, 1950-1, 2 vols.

[27] M. I. Visik, "On general boundary value problems for elliptic partial differential equations," *Trudy Moskovskogo Matematiceskogo Obscestva*, vol. 1 (1952), pp. 187-246.

EIGENFUNCTION EXPANSIONS WITH BOUNDARY CONDITIONS.

By Felix E. Browder.*

Introduction. Let Ω be an open subset of R^n, ρ a positive C^∞ function on Ω. For a fixed integer $r \geqq 1$, let $L^2(\rho)$ be the L^2-space of r-vector functions with respect to the measure $\rho(x)\,dx$ on Ω, (dx = Lebesgue n-measure). $L^2(\rho)$ is a Hilbert space with respect to the inner product

$$(u, v) = \int \langle u(x), v(x) \rangle \rho(x)\,dx \tag{1}$$

where $\langle w_1, w_2 \rangle$ is the natural inner product on C^r. We extend (u, v) to a pairing between u in $C_0^\infty(\Omega)$, the space of C^∞ r-vector functions with compact support in Ω, and v in $\mathcal{D}'(\Omega)$, the space of r-vector distributions on Ω.

Let T be a closed densely defined linear operator in $L^2(\rho)$ which is self-adjoint, i.e. $T = T^*$. If T has a compact resolvent $(T - \lambda I)^{-1}$ for some $\lambda \in C^1$ with $\operatorname{Im}(\lambda) \neq 0$ (and hence for all such λ), then T has a complete orthonormal sequence of eigenfunctions $\{\phi_j\}$ with

$$\phi_j \in \bigcap_{s \geqq 1} D(T^s), \quad T\phi_j = \lambda_j \phi_j, \quad (\lambda_j \in R^1), \tag{2}$$

where $D(T^s)$ is the domain of T^s, and for every u in $L^2(\rho)$

$$u = \sum_{j=1}^{\infty} (u, \phi_j)\phi_j. \tag{3}$$

For more general operators T, there need not exist any proper eigenfunctions in the sense of (2) above, and the general eigenfunction expansion theory asks instead for a decomposition of $L^2(\rho)$ in improper eigenfunctions of T in the following sense:

Definition 1. *By an eigenfunction expansion for T in H is meant three sequences $\{d\sigma_j\}$, $\{\phi_j\}$ and $\{H_j\}$ where:*

1) *The H_j are mutually orthogonal closed subspaces of $L^2(\rho)$ with*

$$L^2(\rho) = \sum_j \oplus H_j. \tag{4}$$

Received June 15, 1964.

* Supported in part by the Sloan Foundation, the Army Research Office (DURHAM) (ARO(D)-31-124-G455), and National Science Foundation grants (G19751 and GP2283).

2) *For each j, $d\sigma_j$ is a finite positive measure on R^1. Let $L^2(d\sigma_j)$ be the L^2-space of complex-valued scalar functions on R^1 with respect to $d\sigma_j$.*

3) *For each j, ϕ_j is a σ_j locally weakly square-integrable function from R^1 to $\mathfrak{D}'(\Omega)$. For each f in $C_o^\infty(\Omega)$, if we set*

$$U_j(f)(\lambda) = (f, \phi_j(\lambda)) \tag{5}$$

then $U_j(f) \in L^2(d\sigma_j)$ and U_j can be extended by continuity to a bounded linear mapping of $L^2(\rho)$ into $L^2(d\sigma_j)$ which maps H_j unitarily onto $L^2(d\sigma_j)$.

4) *The map U of $L^2(\rho)$ into $\sum_j \oplus L^2(d\sigma_j)$ given by*

$$(Uu)_j = U_j(u)$$

is a unitary mapping onto $\sum_j \oplus L^2(d\sigma_j)$.

5) *Let $V_j\colon L^2(\sigma_j) \to H_j$ be given by $V_j = U_j^{-1}$. Then for every $g \in L_o^2(d\sigma_j)$, the set of functions with bounded supports,*

$$V_j(g) = \int_{R^1} g(\lambda)\phi_j(\lambda)\, d\sigma_j(\lambda), \tag{6}$$

the integral converging in $\mathfrak{D}'(\Omega)$.

6) *For each Borel set S in R^1 each f in $L^2(\rho)$, and each j*

$$U_j(E(S)f)(\lambda) = \mathfrak{X}_S(\lambda)U_j(f)(\lambda) \tag{7}$$

where $\mathfrak{X}_S$ is the characteristic function of S.

7) *For $f \in D(T)$,*

$$U_j(Tf)(\lambda) = \lambda U_j(f)(\lambda). \tag{8}$$

If $f \in L^2(\rho)$, then $f \in D(T)$ if and only if

$$\lambda U_j(f) \in L^2(d\sigma_j), \sum_j \| \lambda U_j f \|^2_{L^2(d\sigma_j)} < +\infty.$$

In particular, if f and Tf both lie in $C_o^\infty(\Omega)$, then

$$(\phi_j(\lambda), Tf) = \lambda(\phi_j(\lambda), f), \quad (\lambda \in R^1). \tag{9}$$

The general eigenfunction expansion theorem then asserts:

THEOREM 1. *Every self-adjoint operator T in $L^2(\rho)$ has an eigenfunction expansion in the sense of Definition* 1.

In particular, let A be a system of r-linear partial operators acting on

r-vector functions on Ω, A having coefficients in $C^\infty(\Omega)$ and such that A is formally self-adjoint in $L^2(\rho)$, i.e.

$$(Au, v) = (u, Av)$$

for all u and v in $C_0^\infty(\Omega)$. If T is a self-adjoint realization of A in $L^2(\rho)$ (i.e. $C_0^\infty(\Omega) \subset D(T)$, $Tu = Au$ for $u \in C_0^\infty(\Omega)$), then T has an eigenfunction expansion in the sense of Definition 1, and from equation (9), it follows that

$$A\phi_j(\lambda) = \lambda\phi_j(\lambda), \quad \lambda \in R^1.$$

Neither the general eigenfunction expansion theorem as given above nor its sharpened forms in the literature for elliptic and hypoelliptic operators give any information on the smoothness of the eigenfunctions at boundary points nor on their satisfying boundary conditions imposed on the elements of $D(T)$.

It is our object in the present paper to develop a sharpened form of the general eigenfunction expansion theorem which does deal with boundary smoothness and boundary conditions on the improper eigenfunctions $\phi_j(\lambda)$ and on the related kernels $e_S(x, y)$ of the spectral family $\{E_\lambda\}$. For elliptic differential boundary value problems, partial results especially on the spectral function has been announced in a recent note of Berezanski [6] who derives them however from sharp regularity results for the solutions of elliptic local boundary value problems. Our more general argument applies immediately in the broader context of non-local elliptic boundary value problems on unbounded domains along the lines treated by the writer in [15], [16], [17] (see also Beals [2]).

To formulate our results precisely, we consider the following Assumption and Definition:

ASSUMPTION. *We assume that Γ_0 is an open subset of the boundary Γ of Ω, that Γ_0 is locally a C^∞-manifold with a C^∞ imbedding in R^n, and that $\rho(x)$ is C^0 on an open set Ω' of R^n which contains $\Omega \cup \Gamma_0$.*

For $q \geqq 0$, let

$$C^q(\Omega \cup \Gamma_0) = \{u \mid u \in C^q(\Omega)\,;\ \text{For } |\alpha| \leqq q,\ D^\alpha u \text{ is uniformly continuous on a neighborhood in } \Omega \text{ of each } x_0 \text{ of } \Gamma_0.\}$$

DEFINITION 2. *If q is a non-negative integer, T is said to be q-regular over $\Omega \cup \Gamma_0$ if $\bigcap_{s \geqq 1} D(T^s)$ is contained in $C^q(\Omega \cup \Gamma_0)$.*

THEOREM 2. *Let T be a self-adjoint operator in $L^2(\rho)$ which is q-regular over $\Omega \cup \Gamma_0$ for a given $q \geqq 0$. Then:*

a) *For every bounded Borel set S of R^1, $E(S)$ is an integral operator on $L^2(\rho)$ with*

$$\{E(S)f\}(x) = \int_\Omega e_S(x,y)f(y)\rho(y)\,dy$$

where the kernel e_S is C^q over $(\Omega \cup \Gamma_0) \times (\Omega \cup \Gamma_0)$ while for every compact subset K of $\Omega \cup \Gamma_0$, $\sup_{x \in K} \int |D_x{}^\alpha e_S(x,y)|^2 \rho(y)\,dy < +\infty$, $|\alpha| \leqq q$.

b) *If γ is a continuous linear mapping of $C^q(\Omega \cup \Gamma_0)$ into a Banach space such that $\gamma u = 0$ for all u in $\bigcap_{s \geqq 1} D(T^s)$, then for each y in $\Omega \cup \Gamma_0$, $\gamma\{e_S(\cdot, y)\} = 0$.*

c) *If $\{\phi_j\}$ are the eigenfunctions of the eigenfunction expansion for T of Theorem 1, then ϕ_j may be taken as C^q functions from $\Omega \cup \Gamma_0$ to the weak topology on $L^2(d\sigma_j)$, while $\phi_j(x,\lambda)$ is $(dx \times d\sigma_j)$-measurable on $\Omega \times R^1$.*

THEOREM 3. *Suppose that T is a self-adjoint operator in $L^2(\rho)$ which is $(q+[n/2]+1)$-regular over $\Omega \cup \Gamma_0$ for a given $q \geqq 0$. ($[s]$ is the greatest integer $\leqq s$.) Then:*

a) *For any eigenfunction expansion as in Theorem 1, the eigenfunctions $\{\phi_j\}$ may be taken as σ_j-measurable functions from R^1 to $C^q(\Omega \cup \Gamma_0)$ such that for compact $K \subset \Omega \cup \Gamma_0$, and each compact K_1 in R^1,*

$$\int_{K_1} \{ \sum_{|\alpha| \leqq q} \sup_{x \in K} |D^\alpha \phi_j(x,\lambda)|^2 \} d\sigma_j(\lambda) < +\infty.$$

b) *For λ in the complement of a σ_j-null set, $\phi_j(\cdot,\lambda)$ is the limit in $C^q(\Omega \cup \Gamma_0)$ of a sequence from $\bigcap_{s \geqq 1} D(T^s)$. Hence if $\gamma u = 0$ for all u in $\bigcap_{s \geqq 1} D(T^s)$ where γ is continuous on $C^q(\Omega \cup \Gamma_0)$, then*

$$\gamma(\phi_j(\cdot,\lambda)) = 0.$$

THEOREM 4. *Suppose that T is a self-adjoint operator in $L^2(\rho)$ and that for $k \leqq h + [n/2] + 1$, we know that if $u \in D(T)$ and if u,*

$$Tu \in W_{loc}{}^{k-\delta,2}(\Omega \cup \Gamma_0),$$

for a given $\delta > 0$ then $u \in W_{loc}{}^{k,2}(\Omega \cup \Gamma_0)$ where

$$W_{loc}{}^{k,2}(\Omega \cup \Gamma_0) = \{u \mid u \in L_{loc}{}^2(\Omega \cup \Gamma_0), D^\alpha u \in L_{loc}{}^2(\Omega \cup \Gamma_0) \text{ for } |\alpha| \leqq k\}.$$

Then T is h-regular over $\Omega \cup \Gamma_0$, and if $h = q$ or $h = q + [n/2] + 1$, the conclusions of Theorems 2 and 3, respectively, follow.

The hypothesis of Theorem 4 will certainly hold for non-local elliptic problems of order m for which we postulate that

$$u \in L^2(\rho), Tu \in W_{loc}^{k,2}(\Omega \cup \Gamma_0) \Rightarrow u \in W_{loc}^{k+m}(\Omega \cup \Gamma_0).$$

We observe that the proof of Theorem 4 is a simple consequence of the Sobolev imbedding theorem ([14], Lemma 5).

Proof of Theorem 4. By the Assumption on Γ_0 and ρ, $u \in L_{loc}^2(\Omega \cup \Gamma_0)$. Proceeding by a finite reduction, if $u \in \bigcap_{s \geqq 1} D(T^s)$, then

$$Tu \in \bigcap_{s \geqq 1} D(T^s),$$

and if

$$u, Tu \in W_{loc}^{k-\delta,2}(\Omega \cup \Gamma_0)$$

then $u \in W_{loc}^{k,2}(\Omega \cup \Gamma_0)$ for $k \leqq h + [n/2] + 1$. Hence

$$\bigcap_{s \geqq 1} D(T^s) \subset W_{loc}^{h+[n/2]+1,2}(\Omega \cup \Gamma_0)$$

while by Sobolev's Imbedding Theorem

$$W_{loc}^{h+[n/2]+1,2}(\Omega \cup \Gamma_0) \subset C^h(\Omega \cup \Gamma_0). \qquad \text{Q.E.D.}$$

The proofs given for the Theorems below are simple and self-contained, making no use of any delicate results from functional analysis. Aside from the Sobolev imbedding theorem, they involve only elementary measure theory and the standard form of the spectral theorem for an unbounded self-adjoint operator in a Hilbert space.

In Section 1, we develop the general framework for the proof of our eigenfunction expansion theorems and give a brief proof of Theorem 1. Section 2 gives the proofs of Theorems 2 and 3.

We note that the first general results on eigenfunction expansions for partial differential operators with non-discrete spectrum were those for the Schroedinger operator due to Carleman [18] and Povsner [37]. (The book of Titchmarsh [2] gives a detailed discussion of this case emphasizing the concrete-analytic methods of Levitan and Titchmarsh.) Abstract eigenfunction expansions were first considered by Mautner [34] for operators T with a resolvent which is an operator of Carleman type. These results were extended and applied to general self-adjoint elliptic operators independently by Browder ([7], [8]) and Gårding ([20], [21]). (An extension to hypoelliptic operators with constant coefficients was given by Hörmander in his thesis [28].) The first general eigenfunction expansion theorem for an

arbitrary self-adjoint operator T, and hence for a partial differential operator of arbitrary type, was given by Gelfand and Kostyucenko [22] (see also [23], [24] and [25]), other proofs given by Berezanski [3] and by Browder ([9], [10]), with an extension given in the latter to symmetric operators. Extensions to non-symmetric operators of various classes (subnormal operators, spectral operators) were given by Browder in [11], [12], and [13]. An extension of Bochner's theorem to general eigenfunction expansions was given by Berezanski [4] and other proofs given by Browder [13] and Maurin ([31], [32]). Other abstract expansion theorems covering the elliptic case were given by Bade and Schwartz [1] (see also [19], vol. 2) and by Nelson [35]. Other proofs of the general eigenfunction expansion theorem have been given by Berezanski [5], G. I. Kac [30], Gelfand and Vilenkin [25], Maurin [33], and Harris [27]. Eigenfunction expansions have been applied to the study of the scattering problem for elliptic operators by Ikebe [29] and Greiner [26]. Finally, an interesting application of the abstract eigenfunction expansion theorem has been given by Odhnoff in [36].

Section 1. Let T be a densely defined self-adjoint linear operator in the separable Hilbert space $H = L^2(\rho)$. By the spectral theorem, there exists a projection-valued measure $E(S)$ on F, the family of Borel subsets S of R^1, such that

$$T = \int_{R^1} \lambda dE_\lambda.$$

If u is an element of H, H_u, the cyclic subspace of H generated by u, is defined to be the smallest closed subspace of H containing $E(S)u$ for all $S \in F$. If $S_1 \in F$, then $E(S_1)H_u \subset H_u$. If $v \in H_u{}^\perp$, the orthogonal complement of H_u in H, then for $S \in F$, $w \in H_u$,

$$(E(S)v, w) = (v, E(S)w) = 0,$$

i. e. $E(S)(H_u{}^\perp) \subset H_u{}^\perp$. Let P_u be the orthogonal projection operator of H on H_u. For any $S \in F$, we have for $v \in H$,

$$(E(S)v, v) = (E(S)P_u v, v) + (E(S)(I - P_u)v, v).$$

By a preceding remark, $E(S)P_u v \in H_u$, $E(S)(I - P_u)v \in H_u{}^\perp$ since $P_u v \in H_u$, $(I - P_u)v \in H_u{}^\perp$. Hence

$$(E(S)P_u v, v) = (E(S)P_u v, P_u v)$$

and

$$(E(S)(I - P_u)v, v) = (E(S)(I - P_u)v, (I - P_u)v) \geqq 0.$$

Thus

$$(E(S)v, v) \geqq (E(S)(P_u v), P_u v).$$

Since $v \in D(T)$ if and only if

$$\int \lambda^2 d(E_\lambda v, v) < +\infty$$

it follows that if $v \in D(T)$, then $P_u v \in D(T)$. In addition $P_u E(S) = E(S) P_u$ for all $S \in F$, since for all $v, w \in H$,

$$\begin{aligned}(E(S)P_u v, w) &= (E(S)P_u v, P_u w) = (P_u v, E(S)P_u w)\\ &= (v, E(S)P_u w) = (P_u E(S)v, w).\end{aligned}$$

We have then for $v \in D(T)$,

$$\begin{aligned}P_u Tv &= P_u \int \lambda dE_\lambda v = \int \lambda d(P_u E_\lambda v)\\ &= \int \lambda d(E_\lambda(P_u v)) = TP_u v,\end{aligned}$$

i. e.

$$TP_u = P_u T.$$

LEMMA (1.1). *There exists a sequence* $\{u_j\}$ *in* H *such that if* $H_j = H_{u_j}$,

$$H = \sum_j \oplus H_j.$$

For each j, *let the Borel measure* σ_j *on* R^1 *be given by* $\sigma_j(S) = (E(S)u_j, u_j)$, *and let* $L^2(d\sigma_j)$ *be the corresponding* L^2*-space of scalar functions on* R^1. *For each simple function* f *in* $L^2(d\sigma_j)$,

$$f = \sum_{k=1}^{r} c_k \mathfrak{X}_{S_k}$$

where $\mathfrak{X}_S$ *is the characteristic function of* $S \in F$, *let*

$$V_j(f) = \sum_{k=1}^{r} c_k E(S_k) u_k.$$

Then:

a) V_j *can be extended by continuity to a unitary map of* $L^2(d\sigma_j)$ *onto* H_j.

b) *For each* j, *let* V_j^{-1} *be the inverse map to* V_j *which maps* H_j *onto* $L^2(d\sigma_j)$ *and set*

$$U_j = V_j^{-1} \circ P_{u_j}$$

$U_j\colon H \to L^2(d\sigma_j)$. *If* $U\colon H \to \sum_j \oplus L^2(d\sigma_j)$ *is given by* $(Uv)_j = U_j v$, *then* U *is unitary.*

c) *For* $v \in H$, $S \in F$

$$U_j(E(S)v(\lambda) = \mathfrak{X}_S(\lambda) U_j(v)(\lambda).$$

d) *If $v \in H$, then v lies in $D(T)$ if and only if*

$$\lambda U_j(v)(\lambda) \in L^2(d\sigma_j) \text{ for each } j$$

and

$$\sum_{j=1}^{\infty} \int_{R^1} |\lambda|^2 |U_j(v)(\lambda)|^2 d\sigma_j(\lambda) < +\infty.$$

Then

$$U_j(Tv)(\lambda) = \lambda U_j(v)(\lambda); \ (\lambda \in R^1, j \geqq 1).$$

Proof of Lemma (1.1). If $f = \sum_k c_k \mathfrak{X}_{S_k}$ with $S_t \cap S_h = \emptyset$ for $t \neq h$, then for any u in H, if

$$V_u(f) = \sum_k c_k E(S_k) u,$$

we have

$$\| V_u(f) \|^2 = \sum_k c_t \bar{c}_k (E(S_t)u, E(S_k)u)$$
$$= \sum_k |c_k|^2 (E(S_k)u, u) = \int_{R^1} |f(\lambda)|^2 d\sigma_u(\lambda).$$

We choose a sequence $\{u_j\}$ by induction so that if $H_j = H_{u_j}$, then

$$H = \sum_j \oplus H_j.$$

Then, if $V_j = X_{u_j}$, it follows from the above that V_j is extensible to a unitary mapping of $L^2(d\sigma_j)$ onto H_j. Moreover, for $v \in H$,

$$\| v \|^2 = \sum_j \| P_j v \|^2 = \sum_j \| V_j^{-1} P_j v \|^2$$
$$= \sum_j \| U_j(v) \|^2.$$

Hence U is a unitary mapping of H onto $\sum_j \oplus L^2(d\sigma_j)$.

If $v_j \in H_j$ and if $v_j = V_j(g)$ for a simple function g, then for any Borel set S,

$$V_j(\mathfrak{X}_S g) = \int_S g(\lambda) dE_\lambda u_j = E(S) \int_{R^1} g(\lambda) dE_\lambda u_j$$
$$= E(S) V_j(g).$$

If we apply V_j^{-1} to both sides and set $v_j = V_j(g)$, we obtain

$$\mathfrak{X}_s g = V_j^{-1}(E(S) v_j)$$

i. e.

$$V_j^{-1}(E(S) v_j) = \mathfrak{X}_s V_j^{-1}(v_j).$$

If $v_j = P_{u_j} v$, we have

$$U_j(E(S)v) = \mathfrak{X}_s U_j(v),$$

and (c) is proved.

Let $v \in H$. Then $v \in D(T)$ if and only if

$$v_j = Pu_j v \in D(T)$$

and

$$\sum_j \| Tv_j \|^2 < +\infty.$$

For each j, $v_j \in D(T)$ if and only if

$$\int |\lambda|^2 d(E_\lambda v_j, v_j) < +\infty.$$

However if $v_j = V_j(g)$, then

$$(E(S)v_j, v_j) = \int_S |g(\lambda)|^2 d\sigma_j(\lambda)$$

Hence

$$|\lambda|^2 d(E_\lambda v_j, v_j) = \int_{R^1} |g(\lambda)|^2 |\lambda|^2 d\sigma_j(\lambda)$$

and $v_j \in D(T)$ if and only if the last integral is finite. Moreover, if $w \in H_j \cap D(T)$, then

$$(Tv_j, w) = \int_{R^1} \lambda d(E_\lambda v, w)$$

where

$$(E(S)v_j, w) = \int_S g(\lambda) V_j^{-1}(w)(\lambda)^* d\sigma_j(\lambda).$$

Hence if $w = V_j(h)$

$$(Tv_j, w) = \int_{R^1} \lambda g(\lambda) \overline{h(\lambda)} d\sigma_j(\lambda)$$

$$= (V_j(\lambda g(\lambda)), w),$$

and

$$Tv_j = V_j(\lambda g(\lambda))$$

i. e.

$$V_j^{-1}(Tv_j)(\lambda) = \lambda V_j^{-1}(v_j)(\lambda).$$

In terms of our original v, we have

$$U_j(Tv)(\lambda) = \lambda U_j(v)(\lambda). \qquad \text{Q. E. D.}$$

On the basis of Lemma (1.1), it is clear that the problem of constructing an eigenfunction expansion satisfying the conditions of Theorems 1, 2, or 3 lies in finding $\phi_j(x, \lambda)$ which will act as kernels for the mappings U_j and V_j. One useful observation in doing so lies in the following simple result.

LEMMA (1.2). *Let $L_c^2(d\sigma_j)$ be the subset of $L^2(d\sigma_j)$ of functions with compact support. Then for f in $L_c^2(d\sigma_j)$, $V_j(f) \in \bigcap_{s \geqq 1} D(T^s)$.*

Proof of Lemma (1.2). If $f \in L_c^2(d\sigma_j)$, then f can be approximated by simple functions f_k in $L^2(d\sigma_j)$ with the supports of f_k all contained in a fixed interval $[-N, +N]$. Then $V_j(f_k) \in \bigcap_{s \geqq 1} D(T^s)$ and

$$\| T^s(f_k - f_t) \| \leqq N^s \| f_k - f_t \|.$$

Hence $T^s f_k$ converges in H as $k \to +\infty$, for each $s \geqq 1$. Since T^s is a closed operator in H, $f \in D(T^s)$ for all $s \geqq 1$. Q. E. D.

To obtain kernel representations of the desired form for the mappings U_j and V_j as well as $E(S)$ with S bounded, we shall use the following five lemmas.

LEMMA (1.3). *Let p be a positive number with $1 < p < +\infty$, $d\sigma$ a finite measure on the measure space M. Suppose that V is a continuous linear mapping of $L^p(d\sigma)$ into $C^0(\Omega \cup \Gamma_0)$, and let $p' = p(p-1)^{-1}$.*

Then there exists a function $k(x,t)$ on $(\Omega \cup \Gamma_0) \times M$ which is measurable with respect to $dx \times d\sigma$ and such that the function $x \to k(x, \cdot)$ is continuous from $\Omega \cup \Gamma_0$ to $L^{p'}(d\sigma)$ with the weak topology with

$$\sup_{x \in K} \int_M | k(x,t) |^{p'} d\sigma(t) < +\infty$$

for every compact subset K of $\Omega \cup \Gamma_0$, while for every f in $L^p(d\sigma)$, $x \in \Omega \cup \Gamma_0$,

$$(Vf)(x) = \int_M k(x,t) f(t)\, d\sigma(t).$$

Proof of Lemma (1.3). For each x in $\Omega \cup \Gamma_0$, $(Vf)(x)$ is a continuous linear functional of f in $L^p(d\sigma)$. Hence there exists an element k_x of $L^{p'}(d\sigma)$ such that

$$(Vf)(x) = \int k_x(t) f(t)\, d\sigma(t).$$

Since for f in $L^p(d\sigma)$

$$\int_M k_x(t) \bar{f}(t)\, d\sigma(t) = (V\bar{f})(x)$$

where $(V\bar{f})(x)$ is continuous in x on $\Omega \cup \Gamma_0$, it follows that the function $x \to k_x$ is continuous from $\Omega \cup \Gamma_0$ to the weak topology on $L^{p'}(d\sigma)$. Furthermore

$$\int_M | k_x(t) |^{p'} d\sigma(t)$$

is continuous on $\Omega \cup \Gamma_0$.

Let us restrict each Vf to a compact subset K of $\Omega \cup \Gamma_0$, where we may assume that K is the closure of its interior. Then all the hypotheses of the Lemma hold with Ω and Γ_0 replaced by the interior of K and its boundary. If k'_x is the family of elements in $L^{p'}(d\sigma)$ obtained for x in K, then $k'_x = k_x$ for all x in K. To show the measurability of $k_x(t)$ in both variables, it suffices to prove it for $k'_x(t)$. Hence we may assume without loss of generality that $\Omega \cup \Gamma_0$ is compact.

To complete the proof of our Lemma, we must show that often a suitable choice of a representative function $k_x(t)$ for each element k_x in $L^{p'}(d\sigma)$, then $k(x,t) = k_x(t)$ is $(dx \times d\sigma)$-measurable on $\Omega \times M$. It suffices to show that there exists such a measurable function $k'(x,t)$ such that $k(x,\cdot) = k'(x,\cdot)$, on the complement of a set of measure zero in Ω. We construct such a function k' as the limit of an approximating sequence $k_\epsilon(x,t)$ where $k_\epsilon(x,t)$ is defined as follows: For $\epsilon > 0$, we choose a continuous partition of unity on $\Omega \cup \Gamma_0$ with diameter $< \epsilon$, i.e. a finite family of non-negative functions $\{\eta_j\}$ from $C_c{}^0(R^n)$ such that $\sum_{j=1}^{r} \eta_j(x) = 1$ for $x \in \Omega \cup \Gamma_0$ while for each j the diameter of $\operatorname{supp}(\eta_j)$ is less than ϵ. For each j, we choose a point x_j in $\operatorname{supp}(\eta_j)$ and set

$$k_\epsilon(x,t) = \sum_{j=1}^{r} \eta_j(x) k(x_j,t).$$

We note that for each x in $\Omega \cup \Gamma_0$, we have

$$\int_M |k_\epsilon(x,t)|^{p'} d\sigma(t) \leqq c$$

for a constant $c > 0$ independent of ϵ and of x. Since $\Omega \cup \Gamma_0$ is compact, it follows that

$$\int_\Omega \int_M |k_\epsilon(x,t)|^{p'} d\sigma(t)\, dx \leqq c_1$$

c_1 independent of ϵ. By construction, each k_ϵ is measurable on $\Omega \times M$. If $f \in L^p(d\sigma)$ and $\phi \in L^p(\Omega)$, then

$$\int_{M\times\Omega} k_\epsilon(x,t) f(t) \phi(x)\, d\sigma(t)\, dx$$

$$= \sum_{j=1}^{r} \int_M \int_\Omega \eta_j(x) k(x_j,.t) f(t) \phi(x)\, d\sigma(t)\, dx = \sum_{j=1}^{r} \int_\Omega (Vf)(x_j) \eta_j(x) \phi(x)\, dx.$$

On $\operatorname{supp}(\eta_j) \cap \Omega$,

$$|(Vf)(x_j) - (Vf)(x)| \leqq \delta(\epsilon, f)$$

where $\delta(\epsilon, f) \to 0$ as $\epsilon \to 0$. Hence as $\epsilon \to 0$

$$\sum_{j=1}^{r} \int_{\Omega} (Vf)(x_j)\eta_j(x)\phi(x)\,dx \to \int (Vf)(x)\phi(x)\,dx.$$

Since the set $\{k_\epsilon\}$ is bounded in $L^{p'}(dx \times d\sigma)$, we may find a sequence $\{\epsilon_r\}$ with $\epsilon_r > 0$ for each r such that $\epsilon_r \to 0$ as $r \to +\infty$ and such that the corresponding kernels $k_r = k_{\epsilon_r}$ converge weakly in $L^{p'}(dx \times d\sigma)$ to a function k' in $L^{p'}(dx \times d\sigma)$. By the weak convergence in $L^{p'}(dx \times d\sigma)$, for each $f \in L^p(d\sigma)$, $\phi \in L^p(dx)$, we have as $r \to +\infty$,

$$\int_{\Omega \times M} k_r(x,t) f(t) \phi(x)\, d\sigma(t)\, dx \to \int_{\Omega \times M} k'(x,t) f(t)\phi(x)\, d\sigma(t)\, dx.$$

By a previous remark, the same sequence of integrals has

$$\int_{\Omega} (Vf)(x)\phi(x)\,dx$$

as its limit. Hence by Fubini's theorem, if

$$(V'f)(x) = \int_M k'(x,t) f(t)\, d\sigma(t),$$

then $(V'f)(x)$ exists for almost all x, $V'f \in L^{p'}(dx)$, and

$$\int_{\Omega} (Vf(x) - V'f(x)]\phi(x)\,dx = 0$$

for all $\phi \in L^p(dx)$. Hence for each f in $L^p(d\sigma)$, there exists a dx-null set N_f in $\bar{\Omega}$ such that $(Vf)(x) = (V'f)(x)$ for $x \in \bar{\Omega} - N_f$, i. e.

$$\int k(x,t) f(t)\, d\sigma(t) = \int k'(x,t) f(t)\, d\sigma(t).$$

Let $\{f_r\}$ be a dense denumerable set in $L^p(d\sigma)$, $N = \bigcup_{fs} N_{f_s}$. Then for $x \in \bar{\Omega} - N$, we have for all s,

$$\int k(x,t) f_s(t)\, d\sigma(t) = \int k'(x,t) f_s(t)\, d\sigma(t)$$

and since $k(x,\cdot)$ and $k'(x,\cdot)$ both belong to $L^{p'}(d\sigma)$, it follows that for each such x, there exists a $d\sigma$-null set M_x such that

$$k(x,t) = k'(x,t), \quad t \in M - M_x.$$

Since $k'(x,t)$ is measurable on $\Omega \times M$, the proof of the Lemma is complete.

Q. E. D.

Remark. Lemma (1.3) is a simple variant of the Dunford-Pettis Theorem [19] in which C^0 replaces L^∞. The proof is considerably less technical for the case considered here, and though by applying the Dunford-Pettis Theorem instead of Lemma (1.3), we could weaken all our hypotheses from C^j to $W^{j,\infty}$, there is little practical value in doing so since the only known methods for proving that $\bigcap_{s\geqq 1} D(T^s)$ lies in $W_{loc}{}^{j,\infty}(\Omega \cup \Gamma_0)$ rest on proving that it is contained in $W_{loc}{}^{k,p}(\Omega \cup \Gamma_0)$ for k or p sufficiently large, $p < +\infty$. However, by applying the Sobolev imbedding theorem, we can never obtain the conclusion that $u \in W_{loc}{}^{j,\infty}(\Omega \cup \Gamma_0)$ without automatically having u in $C^j(\Omega \cup \Gamma_0)$.

LEMMA (1.4). *Suppose, in addition to the hypotheses of Lemma* (1.3), *that V maps $L^p(d\sigma)$ into $C^q(\Omega \cup \Gamma_0)$. Then $x \to k(x, \cdot)$ is a C^q-map of $\Omega \cup \Gamma_0$ into the weak topology on $L^{p'}(d\sigma)$ with*

$$\sum_{|\alpha|\leqq q} \sup_{x\in K} \int_M | D_x{}^\alpha k(x,t)|^{p'} d\sigma(t) < +\infty$$

for each compact subset K of $\Omega \cup \Gamma_0$.

Proof of Lemma (1.4). Let $|\alpha| \leqq q$. Then by the closed graph theorem

$$V_\alpha : f \to D^\alpha\{V(f)\}(x)$$

is a mapping of $L^p(d\sigma)$ into $C^0(\Omega \cup \Gamma_0)$ which satisfies the conditions of Lemma (1.3). Hence there exists a measurable function $k_\alpha(x,t)$ on $(\Omega \cup \Gamma_0) \times M$ such that

$$\sup_{x\in K} \int | k_\alpha(x,t)|^{p'} d\sigma(t) < +\infty$$

for every compact subset K of $\Omega \cup \Gamma_0$ while for every f in $L^p(d\sigma)$ and every x in $\Omega \cup \Gamma_0$

$$D^\alpha\{V(f)\}(x) = \int_M k_\alpha(x,t) f(t)\, d\sigma(t).$$

In addition, the map $x \to k_\alpha(x, \cdot)$ is a continuous map from $\Omega \cup \Gamma_0$ to the weak topology in $L^{p'}(d\sigma)$.

To complete the proof of Lemma (1.4), we must show that the map $x \to k(x, \cdot)$ of Ω into $L^{p'}(d\sigma)$ has $k_\alpha(x, \cdot)$ as the corresponding derivative.

It suffices to show that for every function ϕ in $C_0^\infty(\Omega)$ and every f in $L^p(d\sigma)$ we have

$$\int_M \int_\Omega k(x,t)(D^\alpha\phi)(x)f(t)\,dx d\sigma(t)$$
$$= (-1)^{|\alpha|}\int_M \int_\Omega k_\alpha(x,t)\phi(x)f(t)\,dx d\sigma(t).$$

However, we know that the first term is equal to

$$\int_\Omega \{\int_M k(x,t)f(t)\,d\sigma(t)\})(D^\alpha\phi)(x)\,dx$$
$$= \int_\Omega (Vf)(x)(D^\alpha\phi)(x)\,dx = (-1)^{|\alpha|}\int_\Omega D^\alpha\{V(f)\}(x)\phi(x)\,dx$$
$$= (-1)^{|\alpha|}\int_\Omega \{\int_M k_\alpha(x,t)f(t)\,d\sigma(x)\}\phi(x)\,dx$$
$$= (-1)^{|\alpha|}\int_M \int_\Omega k_\alpha(x,t)f(t)\phi(x)\,dx d\sigma(t).$$

Q. E. D.

LEMMA (1.5). *Suppose, in addition to the hypotheses of Lemma* (1.3), *that* V *maps* $L^p(d\sigma)$ *into* $C^{q+[n/p']+1}(\Omega \cup \Gamma_0)$ (*where* $[s]$ *is the greatest integer* $\leqq s$). *Then the kernel* $k(x,t)$ *of Lemma* (1.3) *may be chosen so that for almost all* t *in* M, $k(\cdot,t)$ *lies* $C^q(\Omega\cup\Gamma_0)$ *and for every compact subset* K *of* $\Omega\cup\Gamma_0$,

$$\int_M \{\sum_{|\alpha|\leqq q} \sup_{x\in K} |D_x^\alpha(k(x,t)|\}^{p'}\,d\sigma(t) < +\infty.$$

This is also true if $q = +\infty$.

Proof of Lemma (1.5). Under the hypotheses of Lemma (1.5), it follows from Lemma (1.4) that the mapping $x \to k(x,\cdot)$ is a $C^{q+[n/p']+1}$ map from $\Omega\cup\Gamma_0$ to the weak topology on $L^{p'}(d\sigma)$ and that for $|\alpha| \leqq q + [n/p'] + 1$, there exists a measurable function $k_\alpha(x,t)$ on $\Omega\times M$ such that for all $\phi\in C_0^\infty(\Omega)$ and all f in $L^p(d\sigma)$,

$$\int_M \int_\Omega k(x,t)(D^\alpha\phi)(x)f(t)\,dx d\sigma(t)$$
$$= \int_M \int_\Omega (-1)^{|\alpha|} k_\alpha(x,t)\phi(x)f(t)\,dx d\sigma(t)$$

while for every compact subset K of $\Omega \cup \Gamma_0$,

$$\sup_{x \in K} \int_M | k_\alpha(x,t) |^{p'} d\sigma(t) < +\infty.$$

We note first that for almost all t in M, $k_\alpha(\cdot, t)$ lies in $L^{p'}(\Omega)$, $|\alpha| \leqq q$. We show next for such α, $D^\alpha k(\cdot, t) = k_\alpha(\cdot, t)$ in the distribution sense for almost all t in M, i. e. for all $\phi \in C_c^\infty(\Omega)$

$$\int_\Omega k(x,t)(D^\alpha \phi)(x)\,dx = (-1)^{|\alpha|} \int_\Omega k_\alpha(x,t)\phi(x)\,dx,$$

To establish this fact, we remark that if $f \in L^p(d\sigma)$ and $\zeta(t)$ denotes the difference of the two integrals in the last equation, then by the equality above and Fubini's theorem,

$$\int \zeta(t) f(t)\, d\sigma(t) = 0.$$

Since $\zeta \in L^{p'}(d\sigma)$, it follows that $\zeta(t) = 0$ for t outside a σ-null set. Hence $D^\alpha k(\cdot, t) = k_\alpha(\cdot, t)$ for t outside the same σ-null set.

Finally outside a σ-null set for t, $k(\cdot, t) \in W_{loc}^{q+[n/p']+1}(\Omega \cup \Gamma_0)$. It follows from the Sobolev imbedding theorem that $k(\cdot, t)$ for such t, lies in $C^q(\Omega \cup \Gamma_0)$, This argument is valid if $q = +\infty$. Q. E. D.

LEMMA (1.6). *The conclusions of Lemmas* (1.2), (1.3), *and* (1.4) *remain valid if $L^p(\sigma)$ is the space of r_1-vector functions on M and if the measure σ is σ-finite rather than finite. In this case the kernel $k(x,t)$ is an $(r \times r_1)$ matrix function on $\Omega \times M$.*

The proof of Lemma (1.6) is straightforward and is omitted here.

LEMMA (1.7). *Let V be a continuous linear mapping of $L^p(d\sigma)$ for a σ-finite measure $d\sigma$ on a measure space M into $L^p(\Omega)$. Then there exists a weakly measurable function k from M to the $(r \times r_1)$-matrices with components in $\mathcal{D}'(\Omega)$ such that*

$$Vf = \int k(t) f(t)\, d\sigma(t)$$

i. e. for every function $\phi \in C_c^\infty(\Omega)$

$$(Vf, \phi) = \int (k(t) f(t), \phi)\, d\sigma(t).$$

For every $\phi \in C_c^\infty(\Omega)$, $(k(t)f(t), \phi)$ is locally summable with respect to $d\sigma$.

Proof of Lemma (1.7). We may assume without loss of generality that Ω is bounded and that σ is a finite measure. We consider $L^p(\Omega)$ imbedded in $L^p(R^n)$ by setting each u in $L^p(\Omega)$ to be zero outside of Ω.

For $2m > n$, let

$$s(x) = (2\pi)^{-n/2} \int_{R^n} \frac{e^{-i\langle x,\xi\rangle} d\xi}{|\xi|^{2m}+1}$$

Then s is a uniformly bounded continuous function on R^n and for every scalar testing function ϕ in $C_c^\infty(R^n)$

$$(s, (-\Delta)^m\phi + \phi) = \int_{R^n} \hat{s}(\xi) \cdot (|\xi|^{2m}+1)(\hat{\phi}(\xi))^* d\xi$$

where $\hat{s}$ and $\hat{\phi}$ are the Fourier transforms of the L^2 functions s and ϕ. However,

$$\hat{s}(\xi) = \frac{1}{|\xi|^{2m}+1}$$

Hence

$$(s, (-\Delta)^m\phi + \phi) = \int_{R^n} (\hat{\phi}(\xi))^* d\xi = \phi^*(0).$$

From the given mapping V, we generate a new mapping V_1 of $L^p(d\sigma)$ into $C^0(R^n)$ by setting

$$V_1 f = s * Vf,$$

where Vf is identified with the corresponding element of $L^p(R^n)$. Since Vf lies in $L^1(R^n)$, $V_1 f$ lies in $C^0(R^n)$ and the map $f \to V_1 f$ is continuous from $L^p(d\sigma)$ to $C^0(R^n)$.

Applying Lemma (1.3), there exists a kernel $k(x,t)$ such that

$$V_1 f(x) = \int k_1(x,t) f(t)\, d\sigma(t).$$

Since $Vf = \{(-\Delta)^m + 1\} V_1 f$, we have

$$Vf = \int k(t) f(t)\, d\sigma(t)$$

where

$$k(t) = \{(-\Delta)^m + 1\} k_1(\cdot, t). \qquad \text{Q.E.D.}$$

Proof of Theorem 1. The mapping V_j of $L^2(d\sigma_j)$ into $H = L^2(\sigma)$ and U_j of H into $L^2(d\sigma_j)$ are adjoint to one another. To show that for $u \in C_c^\infty(\Omega)$

$$U_j(u)(\lambda) = (u, \phi_j(\lambda))$$

and that for f in $L^2(d\sigma_j)$ with bounded support

$$V_j(f) = \int \phi_j(\lambda) d(\lambda) d\sigma_j(\lambda)$$

are equivalent problems since

$$\int U_j(u)(\lambda) f(\lambda)^* d\sigma_j(\lambda) = (u, V_j(f))$$

for such a pair u and f. For V_j, we apply Lemma (1.7). Q.E.D.

Section 2. The present section is devoted to the proofs of Theorems 2 and 3, which are the basic results of our discussion.

Proof of Theorem 2. Proof of (a). If S is a bounded Borel set, $E(S)$ maps $L^2(\rho)$ into $\bigcap_{s \geqq 1} D(T^s)$. If T is q-regular, it follows from Lemma (1.6) that $E(S)$ has a kernel representation of the form

$$\{E(S)f\}(x) = \int_S e_S(x, y) f(y) \rho(y) dy$$

where the function $x \to e_S(x, \cdot)$ is C^q from $\Omega \cup \Gamma_0$ to $L^2(\rho)$ so that for any compact subset K of $\Omega \cup \Gamma_0$, and for $|\alpha| \leqq q$

$$\sup_{x \in K} \int |D_x^\alpha e_S(x, y)|^2 \rho(y) dy < +\infty.$$

Moreover since

$$(E(S)f, g) = (f, E(S)g),$$

we have

$$\int_\Omega \int_\Omega \langle e_S(x, y) f(y), g(x) \rangle \rho(y) \rho(x) dx dy$$
$$= \int \int \langle f(y), e_S(y, x) g(x) \rangle \rho(y) \rho(x) dx dy.$$

Hence it is true a.e. on $\Omega \times \Omega$ that

$$e_S(x, y) = e_S(y, x)^*$$

so that

$$\sup_{y \in K} \int |D_y^\alpha e_S(x, y)|^2 \rho(x) dx < \infty, \quad |\alpha| \leqq q.$$

Finally

$$E(S)^2 = E(S)$$

so that

$$e_S(x,y) = \int e_S(x,z)\cdot e_S(z,y)\rho(z)\,dz.$$

Hence

$$D_x^\alpha D_y^\beta e_S(x,y) = \int (D_x^\alpha e_S)(x,z)\cdot (D_y^\beta e_S)(z,y)\rho(z)\,dz$$

and

$$D_x^\alpha D_y^\beta e_S(x,y) \text{ is continuous on } (\Omega\cup\Gamma_0)\times(\Omega\cup\Gamma_0)$$

for $|\alpha|, |\beta| \leqq q$.

Proof of (b). For any $f\in L^2(\rho)$,

$$E(S)f = \int_\Omega e_S(\cdot,y)f(y)\rho(y)\,dy$$

lies in $\bigcap_{s\geqq 1} D(T^s)$.

Let y_0 be a point of $\Omega\cup\Gamma_0$. Choose a sequence of functions f_k with the support of f_k in the ball of diameter $1/k$ about y_0 such that

$$\int f_k(y)\rho(y)\,dy = 1$$

Let $|\alpha|\leqq q$. Then for $x\in\Omega\cup\Gamma_0$,

$$D^\alpha E(S)f(x) = \int_\Omega e_S^{(\alpha)}(x,y)f(y)\rho(y)\,dy$$

while

$$\int e_S^{(\alpha)}(x,y)f_k(y)\rho(y)\,dy = e_S^{(\alpha)}(x,y_0) + \int [e_S^{(\alpha)}(x,y) - e_S^{(\alpha)}(x,y_0)]f_k(y)\,dy$$

The second integral is bounded in C^0-norm by

$$\max_{|y-y_0|<1/k} |e_S^{(\alpha)}(x,y) - e_S^{(\alpha)}(x,y_0)| \to 0$$

as $k\to\infty$. Hence $E(S)f_k\to e_S(x,y_0)$ in $C^q(\Omega\cup\Gamma_0)$. However, $E(S)f_k$ lies for each k in $\bigcap_{s\geqq 1} D(T^s)$.

Proof of (c). We consider the mapping V_j of $L^2(d\sigma_j)$ into $L^2(\rho)$ and restrict it to $L_K^2(d\sigma_j)$, the subspace of functions having support in a compact subset K of R_1. By Lemma (1.2), V_jf for $f\in L_K^2(d\sigma_j)$ lies in $\bigcap_{s\geqq 1} D(T^s)$, which in turn is contained in $C^q(\Omega\cup\Gamma_0)$. V_j is continuous from $L_K^2(d\sigma_j)$

to $L^2(\rho)$ and hence a closed linear mapping from $L_K^2(d\sigma_j)$ to $C^q(\Omega \cup \Gamma_0)$. Applying the closed graph theorem, V_j is continuous between this last pair of spaces. Applying Lemmas (1.3) and (1.4), the conclusion of part (c) follows. Q. E. D.

Proof of Theorem 3. Part (a) follows by applying Lemma (1.5) to the mapping V_j taken from $L_K^2(d\sigma_j)$ to $C^{q+[n/2]+1}(\Omega \cup \Gamma_0)$. Part (b) follows from part (a) by the same argument as given for part (b) of Theorem 2. Q. E. D.

REFERENCES.

[1] W. Bade and J. T. Schwartz, "On abstract eigenfunction expansions," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 42 (1956), pp. 519-525.

[2] R. W. Beals, "Non-local boundary value problems for elliptic operators," *American Journal of Mathematics*, vol. 87 (1965), pp. 315-362.

[3] Yu. M. Berezanski, "On expansions in eigenfunctions of self-adjoint operators," *Matematičskiĭ Sbornik*, vol. 43 (1957), pp. 75-126.

[4] ———, "Representation of positive definite kernels in terms of eigenfunctions of differential equations," *Matematičeskiĭ Sbornik*, vol. 47 (1959), pp. 145-176.

[5] ———, "On expansions in eigenfunctions of self-adjoint operators," *Ukranskiĭ Matematičeskiĭ Žurnal*, vol. 11 (1959), pp. 16-24.

[6] ———, "On the smoothness of the spectral function of a self-adjoint elliptic differential operator up to the boundary of the domain," *Dokladi Akademii Nauk SSSR*, vol. 52 (1963), pp. 511-514.

[7] F. E. Browder, "The eigenfunction expansion theorem for the general self-adjoint singular elliptic partial differential operator, I." The analytical foundation, *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 40 (1954), pp. 454-459.

]8[———, "Eigenfunction expansions for singular elliptic operators, II. The Hilbert space argument: parabolic equations on open manifolds," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 40 (1954), pp. 459-463.

[9] ———, "Eigenfunction expansions for formally self-adjoint partial differential operators, I," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 42 (1956), pp. 769-771.

[10] ———, "Eigenfunction expansions for formally self-adjoint partial differential operator, II," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 42 (1956), pp. 870-872.

[11] ———, "Eigenfunction expansions for non-symmetric partial differential operators, I," *American Journal of Mathematics*, vol. 80 (1958), pp. 365-381.

[12] ———, "Eigenfunction expansions for non-symmetric partial differential operators, II," *American Journal of Mathematics*, vol. 81 (1959), pp. 1-22.

[13] ———, "Eigenfunction expansions for non-symmetric partial differential operators, III," *American Journal of Mathematics*, vol. 81 (1959), pp. 715-734.

[14] ———, "On the spectral theory of elliptic differential operators, I," *Mathematische Annalen*, vol. 142 (1961), pp. 22-130.

[15] ———, "Families of closed linear operators depending on a parameter," *American Journal of Mathematics (to appear)*.

[16] ———, "Non-local elliptic boundary value problems," *American Journal of Mathematics*, vol. 86 (1964), pp. 735-750.

[17] ———, "Asymptotic distribution of eigenvalues and eigenfunctions for non-local elliptic boundary value problems," *American Journal of Mathematics*, vol. 87 (1865), pp. 175-195.

[18] T. Carleman, "Sur la théorie mathematique d'equation de Schrödinger," *Arkhiv. för Matematik, Astronomi och Fysik*, vol. 24, No. 11 (1934), pp. 1-7.

[19] N. Dunford and J. T. Schwartz, *Linear operators*, 2 vols.. New York, 1956-1963.

[20] L. Gårding, "Eigenfunction expansions connected with elliptic differential operators," *Proceedings, 12th Scandinavian Mathematical Congress*, Lund, (1954), pp. 44-55.

[21] ———, *Applications of the theory of direct integrals of Hilbert spaces to some integral and differential operators*, Lecture Notes No. 11, Institute Fluid Dynamics and Applied Mathematics, University of Maryland (1954).

[22] I. M. Gelfand and A. G. Kostyucenko, "On expansions in characteristic functions of differential and other operators," *Dokladi Akademii Nauk SSSR*, vol. 103 (1955), pp. 349-352.

[23] ——— and G. E. Silov, "Quelques applications de la theorie des fonctions géneralisées," *Journal de Mathématiques Pures Appliquees*, vol. 35 (1956), pp. 383-413.

[24] ——— and ———, *Some questions in the theory of differential equations. (Generalized functions, vol. 3)*, Moscow, 1958.

[25] ——— and N. Ya. Vilenkin, *Some applications in harmonic analysis. Outer Hilbert spaces. (Generalized functions, vol. 4)*, Moscow, 1961.

[26] P. Greiner, *Eigenfunction expansions and scattering theory for perturbed elliptic partial differential operators*, Yale Ph. D. Dissertation, 1964.

[27] R. T. Harris, "Generalized eigenfunction expansions for operator algebras," *Transactions of the American Mathematical Society*, vol. 98 (1961), pp. 485-500.

[28] L. Hörmander, "On the theory of general partial differential operators," *Acta Mathematica*, vol. 94 (1955), pp. 161-248.

[29] T. Ikebe, "Eigenfunction expansions associated with the Schroedinger operators and their application to scattering theory," *Archives for Rational Mechanics and Analysis*, vol. 5 (1960), pp. 1-34.

[30] G. I. Kac, "Expansions in eigenfunctions for self-adjoint operators," *Dokladi Akademii Nauk SSSR*, vol. 119 (1958), pp. 19-22.

[31] K. Maurin, "Entwicklung positiv definiter Kerne nach Eigendistributionen. Differenzbarkeit der Spectralfunktion eines hypermaximalen Operators," *Bulletin de l'Academie Polonaise des Sciences. Série des Sciences Mathematiques*, vol. 6 (1958), pp. 149-155.

[32] ———, "Spektraldarstellung der Kerne Eine Verall gememerung der Sätze von

Källen-Lehman und Herglotz-Bochner," *Bulletin de l'Academie Polonaise des Sciences. Série des Sciences Mathematiques*, vol. 7 (1959), pp. 461-470.

[33] ———, "Allgemeine Eigenfunktionen-entwicklungen. Spektraldarstellung abstrakter Kerne. Eine Verallgemeinerungen der Distributionen auf Lee'-schen Gruppen," *Bulletin de l'Academie Polonaise des Sciences, Série des Sciences Mathematiques*, vol. 7 (1959), pp. 471-479.

[34] F. I. Mautner, "On eigenfunction expansions," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 39 (1953), pp. 49-53.

[35] E. Nelson, "Kernel functions and eigenfunction expansions," *Duke Mathematical Journal*, vol. 25 (1958), pp. 15-27.

[36] J. Odhnoff, *Operators generated by differential problems with eigenvalue parameter in equation and boundary conditions*. Thesis, Lund (1959), pp. 1-80.

[37] A. Ya Povsner, "On the expansion of arbitrary functions in eigenfunctions of the operator $-\Delta u + cu$," *Matematičeskiĭ Sbornik*, vol. 32 (1953), pp. 109-156.

[38] F. Riesz and B. Sz Nagy, *Functional analysis*, New York, 1955.

[39] L. Schwartz, *Théorie des distributions*, 2 vols., Paris, 1950-1951.

[40] ———, "Théorie des noyaux," *Proceedings of the International Congress of Mathematicians*, Cambridge (1952), pp. 220-230.

[41] M. H. Stone, *Linear transformations in Hilbert space and their applications to analysis*.

[42] E. C. Titchmarsh, *Eigenfunction expansions associated with second order differential equations*, Part II, Oxford, 1958.

[43] J. von Neumann, "On rings of operators, Reduction theory," *Annals of Mathematics*, vol. 50 (1949), pp. 401-485.

ON ISOMORPHISMS OF C^*-ALGEBRAS.[1]

By L. TERRELL GARDNER.

Chapter 1.

Introduction. In this paper we study mappings of C^*-algebras which preserve the complex algebraic structure, but not necessarily the $*$-structure of these algebras. The term "isomorphism" will be used to denote such a map, and "$*$-isomorphism" to denote an isomorphism ψ which satifies the additional condition $\psi(A^*) = \psi(A)^*$.

The paper is directed toward the proof of two main theorems, both of which appear in Chapter 4: Theorem A gives the structure of isomorphisms, showing that they are in a certain canonical sense "spatial" in nature. Theorem B asserts that if two C^*-algebras are isomorphic, then they are $*$-isomorphic. S. Sakai has formulated this problem for the case of W^*-algebras [10, p. 1. 53, problem i)].

In Chapter 2 we introduce notation and summarize the necessary representation theory. Chapter 3 studies certain automorphisms of C^*-algebras, culminating in the "Invariance Theorem," which is applied in the following form: If $\mathfrak{A}$ is a C^*-algebra of operators acting on the Hilbert space $\mathcal{H}$, and T is a positive, invertible operator in $\mathfrak{L}(\mathcal{H})$ then $T^{-1}\mathfrak{A}T = \mathfrak{A}$ implies that $T^{-(\frac{1}{2})}\mathfrak{A}T^{\frac{1}{2}} = \mathfrak{A}$, also.

The author wishes here to express his gratitude to Professor Richard V. Kadison, who directed his attention to the problem of Theorem B, and without whose advice and encouragement this work would not have been done.

Chapter 2.

Preliminaries; Representation theory. We begin by remarking that all of our Hilbert spaces $\mathcal{H}$ are over the complex numbers; we put no restriction on the Hilbert dimension of $\mathcal{H}$. All of our Banach algebras are complex, and have unit element. By C^*-algebra we mean an abstract Banach $*$-algebra $\mathfrak{A}$ with $\| A^*A \| = \| A^* \| \| A \|$ for all $A \in \mathfrak{A}$ (B^*-algebra). A representation ($*$-representation) of $\mathfrak{A}$ on the Hilbert space $\mathcal{H}$ is a homo-

Received June 25, 1964.

[1] This work was supported in part by National Science Foundation grants NSF GP-1604 and NSF G-19022.

morphism (*-homomorphism) of $\mathfrak{A}$ into $\mathfrak{L}(\mathcal{H})$, the algebra of all bounded operators on $\mathcal{H}$. A *-representation ϕ on $\mathcal{H}$ is *cyclic* if there exists a vector x in $\mathcal{H}$ (cyclic vector) such that the closure $[\phi(\mathfrak{A})x]$ of $\{\phi(A)x \mid A \in \mathfrak{A}\}$ is $\mathcal{H}$. It is *irreducible* if every $x \neq 0$ in $\mathcal{H}$ is cyclic. Kadison [6] has shown that the topological irreducibility thus defined is equivalent to algebraic irreducibility. A *-representation is *faithful* if it is a *-isomorphism, in which case it is an isometry; it is *fully reducible* if it is a direct sum of irreducible representations.

A classical theorem of Gelfand-Neumark [2], as strengthened and elegantly set forth in [3], asserts that every C^*-algebra has a faithful *-representation as a C^*-algebra of operators on a suitable Hilbert space.

Two *-representations ϕ, ψ of $\mathfrak{A}$ on $\mathcal{H}$, $\mathcal{K}$, respectively, are said to be (unitarily) *equivalent* $(\phi \sim \psi)$ if and only if there is a unitary transformation U of $\mathcal{H}$ onto $\mathcal{K}$ such that $U\phi(A)U^* = \psi(A)$ for all A in $\mathfrak{A}$.

A *state* of $\mathfrak{A}$ is a positive linear functional ρ with $\rho(I) = 1$. The *left kernel*, $\mathfrak{J}_\rho$, of the state ρ is the set of A in $\mathfrak{A}$ such that $\rho(A^*A) = 0$. By the Schwarz inequality for ρ, $\mathfrak{J}_\rho$ is a left ideal. $\mathfrak{A}/\mathfrak{J}_\rho$ is therefore a left $\mathfrak{A}$-module in a natural way, and the algebraic representation of $\mathfrak{A}$ on $\mathfrak{A}/\mathfrak{J}_\rho$ is denoted ϕ_ρ. We define on $\mathfrak{A}/\mathfrak{J}_\rho$ a positive definite inner product $(A + \mathfrak{J}_\rho, B + \mathfrak{J}_\rho) = \rho(B^*A)$, and after verifying that $\phi_\rho(A)$ is bounded for each $A \in \mathfrak{A}$, we extend $\phi_\rho(A)$ to a bounded operator on $\mathcal{H}_\rho = (\mathfrak{A}/\mathfrak{J}_\rho)^-$, the completion of the prehilbert space $\mathfrak{A}/\mathfrak{J}_\rho$ in its norm $\| \cdot \|_\rho$. We call the map thus defined on $\mathfrak{A}$ to $\mathfrak{L}(\mathcal{H}_\rho)$ again ϕ_ρ; ϕ_ρ is a cyclic *-representation of $\mathfrak{A}$ on $(\mathfrak{A}/\mathfrak{J}_\rho)^-$, with cyclic vector $I + \mathfrak{J}_\rho$; it is called the *representation due to* ρ.

We say ρ is a *pure state* of $\mathfrak{A}$ if ρ is an extreme point of the weak-* compact, convex set of states of $\mathfrak{A}$. We denote the set of pure states of $\mathfrak{A}$ by $\mathfrak{P}(\mathfrak{A})$. If and only if ρ is pure, $\mathfrak{J}_\rho$ is a maximal left ideal, ϕ_ρ is irreducible, and $\mathfrak{A}/\mathfrak{J}_\rho$ is complete in the inner-product norm defined above [6].

Every cyclic *-representation of $\mathfrak{A}$ is equivalent to the representation due to some state of $\mathfrak{A}$. If ϕ represents $\mathfrak{A}$ on $\mathcal{H}$ with cyclic vector x, and $\| x \| = 1$, put $\rho(A) = (\phi(A)x, x)$, $A \in \mathfrak{A}$. Then ϕ_ρ and ϕ are equivalent, the relevant unitary mapping being the extension to $(\mathfrak{A}/\mathfrak{J}_\rho)^-$ of the isometry $A + \mathfrak{J}_\rho \to \phi(A)x$ of $\mathfrak{A}/\mathfrak{J}_\rho$ into $\mathcal{H}$. Since irreducibility is obviously a unitary invariant, each irreducible *-representation is equivalent to the representation due to some pure state.

A state ω_x defined on $\mathfrak{L}(\mathcal{H})$ by a unit vector x in $\mathcal{H}$ by the equation $\omega_x(A) = (Ax, x)$ is called a *vector state*. If $\mathfrak{A}$ is a *-subalgebra of $\mathfrak{L}(\mathcal{H})$, we may call $\omega_x |_{\mathfrak{A}}$ a vector state of $\mathfrak{A}$.

We define equivalence of states in terms of their associated representations: $\rho_1 \sim \rho_2$ if and only if $\phi_{\rho_1} \sim \phi_{\rho_2}$. In [3], it is shown that two pure states ρ, μ of $\mathfrak{A}$ are equivalent if and only if there exists a unitary operator U in $\mathfrak{A}$ such that $\mu(A) = \rho(U^*AU)$ for all A in $\mathfrak{A}$. In [6], it is shown that the correspondence between maximal left ideals and pure states is one-one: a maximal left ideal $\mathfrak{J}$ is contained in the null space of a unique pure state ρ, and $\mathfrak{J} = \mathfrak{J}_\rho$. (See also Corollary 2.3 and the proof of Lemma 4.4 in the present paper.)

The *universal representation* of $\mathfrak{A}$ is the direct sum $\Phi = \bigoplus_\rho \phi_\rho$ over all states ρ of $\mathfrak{A}$. Φ is faithful, and is universal in the sense that every cyclic representation of $\mathfrak{A}$ is equivalent to a subrepresentation of Φ, but universal also in the stronger sense: if we identify $\mathfrak{A}$ with its image under Φ and let $\mathfrak{A}^-$ denote its strong closure, then for each $*$-representation ψ of $\mathfrak{A}$, there is a unique weakly-continuous extension of ψ which maps $\mathfrak{A}^-$ $*$-homomorphically onto $\psi(\mathfrak{A})^-$ [4; 7].

THEOREM 2.1 (Sherman [12], Takeda [13]). *Let $\mathfrak{A}$ be identified with its image under its universal representation, and let $\mathfrak{A}^-$ be its strong closure; then $\mathfrak{A}^-$ is isomorphic as a Banach space to the second dual space $\mathfrak{A}''$ of $\mathfrak{A}$. If $\mathfrak{A}$ is considered as canonically embedded in $\mathfrak{A}''$, then the isomorphism can be taken as an extension of the identity map of $\mathfrak{A}$. (For a succinct proof, we recommend [7]; see also [4].)*

Since every state of $\mathfrak{A}$ "is" a vector state in the universal representation, the weak and ultraweak topologies coincide on $\mathfrak{A}^-$. Since the (vector) states linearly span the dual $\mathfrak{A}''$ of the C^*-algebra $\mathfrak{A}$, the weak (weak-operator) and weak-$*$-topologies coincide on $\mathfrak{A}^- = \mathfrak{A}''$, also. If ρ is a state of $\mathfrak{A}$, and $\rho = \omega_x |_{\mathfrak{A}}$ when $\mathfrak{A}$ is taken in its universal representation, put $s(\rho)$ (the *support* of ρ) $= \inf\{E \mid E$ is a projection in $\mathfrak{A}^-$ and $x \in E\} = \inf\{E \mid E$ is a projection in $\mathfrak{A}^-$ and $\omega_x(E) = 1\}$. Clearly $x \in s(\rho)$, and if we continue to denote by ρ its unique weakly continuous extension ($= \omega_x |_{\mathfrak{A}^-}$) to $\mathfrak{A}^-$, we have $\rho(s(\rho)) = 1$.

THEOREM 2.2 [7, Theorem 2.5]. *Let $\mathfrak{A}$ be a C^*-algebra, $\mathfrak{J}$ a norm-closed left ideal in $\mathfrak{A}$. Then there is a projection E in $\mathfrak{A}''$ with the properties*

i) $\mathfrak{J} = \mathfrak{A} \cap \mathfrak{A}''E$;

ii) $\mathfrak{J}^- = \mathfrak{A}''E$.

Moreover, the projection E with property ii) *is uniquely determined by $\mathfrak{J}$. ($\mathfrak{J}^-$ denotes the strong (equivalently weak, or ultraweak) closure of $\mathfrak{J}$ in $\mathfrak{A}^- = \mathfrak{A}''$.)*

Proof. The uniqueness claim is trivial: Clearly $\mathfrak{J}$ determines $\mathfrak{J}^-$; if $\mathfrak{J}^- = \mathfrak{A}''E = \mathfrak{A}''F$, with E and F projections, then $F = FE = F^* = EF = E$. The existence proof employs from [6] the following results:

THEOREM 2.2.1. *$\mathfrak{J}$ is generated as a left ideal by its positive elements.*

THEOREM 2.2.2. *$\mathfrak{J}$ is the intersection of the left kernels of the (pure) states that annihilate it.*

We write, here and later, $n(T)$ for the (projection on the) null space of T, $s(T) = I - n(T)$, $r(T) =$ (projection on the) closure of the range of T. Now let $E = \sup\{r(T) \mid T \in \mathfrak{J}, T \geqq 0\}$. Then each positive T in $\mathfrak{J}$ lies in $\mathfrak{A}''E$, so that by Theorem 2.2.1, $\mathfrak{J} \subset \mathfrak{A}''E$. Since $\mathfrak{A}''$ is strongly closed, $\mathfrak{J}^- \subset \mathfrak{A}''E$. Clearly, $I - E = \sup\{F \mid F$ is a projection in $\mathfrak{A}''$ and $\mathfrak{J}F = 0\}$; that is, $(I - E)\mathcal{H}$ is precisely the annihilator of $\mathfrak{J}$. In the universal representation space $\mathcal{H}$, $x \in (I - E)\mathcal{H}$ if and only if $\omega_x(A) = 0$ for all A in $\mathfrak{J}$. But by Theorem 2.2.2, and the fact that each state ρ of $\mathfrak{A}$ is $\omega_x \mid_{\mathfrak{A}}$ for some $x \in \mathcal{H}$, $\mathfrak{J} = \{A \in \mathfrak{A} \mid \omega_x(A^*A) = 0$ for all x in $(I - E)\mathcal{H}\}$ $= \{A \in \mathfrak{A} \mid \| Ax \|^2 = 0$, all $x \in (I - E)\mathcal{H}\}$, so $\mathfrak{J}$ is the annihilator of $(I - E)\mathcal{H}$ in $\mathfrak{A}$. Now if $E' = \sup\{r(T) \mid T \in \mathfrak{J}^-, T \geqq 0\}$, we have $E' \supset E$, but from $\mathfrak{J}^- \subset \mathfrak{A}''E$, we have $E' \subset E$, so that $E' = E$, $\mathfrak{J}^-$ is the annihilator in $\mathfrak{A}''$ of $(I - E)\mathcal{H}$, and $\mathfrak{J} = \mathfrak{A} \cap \mathfrak{J}^-$. Especially, $E \in \mathfrak{J}^-$, so that $\mathfrak{A}''E \subset \mathfrak{J}^-$, and $\mathfrak{A}''E = \mathfrak{J}^-$. Thus finally, $\mathfrak{A} = \mathfrak{A} \cap \mathfrak{A}''E$.

COROLLARY 2.3. *Let $\mathfrak{A}$ be a C^*-algebra, ρ a pure state of $\mathfrak{A}$, and $E = s(\rho)$; then*

i) $\mathfrak{A}''(I - E) = \mathfrak{J}^-{}_\rho$;

ii) *E is a minimal projection in $\mathfrak{A}''$.*

Proof. It is clear from Theorem 2.2 and its proof that our task is to show that $(I - E) = \sup\{r(T) \mid T \geqq 0$ and $T \in \mathfrak{J}_\rho\}$; that is, with $G = \inf\{n(T) \mid T \geqq 0$ and $T \in \mathfrak{J}_\rho\}$, we wish to show $E = G$.

a) $E \subset G$. This will follow, if $\rho(n(T)) = 1$ for all positive T in $\mathfrak{J}_\rho$. We choose a unit vector x in $\mathcal{H}$, the universal representation space for $\mathfrak{A}$, with the property $\omega_x \mid_{\mathfrak{A}} = \rho$. Then $T \geqq 0$, T in $\mathfrak{J}_\rho$ means $\omega_x(T^2) = 0$, or $Tx = 0$; hence $x \in n(T)$, and $\omega_x(n(T)) = 1$, as desired.

b) $E = G$. If not, put $F = G - E$. The left annihilator $\mathfrak{J}$ of F contains E, as $\mathfrak{J}^-{}_\rho$ does not. But if $A \in \mathfrak{J}_\rho$, $AF = AGF = 0$, since G is, by earlier remarks, the annihilator of $\mathfrak{J}_\rho$. We find that $\mathfrak{J} \supsetneq \mathfrak{J}^-{}_\rho$, contradicting the maximality of $\mathfrak{J}^-{}_\rho$.

The argument above shows G $(=E)$ to be minimal in $\mathfrak{A}''$, as claimed.

From the (easily proved) faithfulness of the universal representation, one passes by way of the Krein-Milman theorem to the fact that $\Gamma = \bigoplus_{\rho \in \mathcal{P}(\mathfrak{A})} \phi_\rho$ is also faithful.

The following propositions are known. Our proofs simply indicate their relation to certain of the foregoing results.

PROPOSITION 2.4. *The projection lattice of* $\Gamma(\mathfrak{A})^-$ *is purely atomic.*

Proof. We need only show that if E is a non-zero projection in $\Gamma(\mathfrak{A})^-$, then $E \supset F$ for some minimal projection F in $\Gamma(\mathfrak{A})^-$. Choose x, $0 \neq x \in E$; then $x = \bigoplus_\rho x_\rho$, $x_\rho \in \mathcal{H}_\rho$. Put $z = x_\rho / \| x_\rho \|$ for some fixed ρ with $\| x_\rho \| \neq 0$. Then ω_z is a vector state. $\omega_z \mid \phi_\rho(\mathfrak{A})$ is a vector state of the irreducible algebra $\phi_\rho(\mathfrak{A})$ on $\mathcal{H}_\rho$, hence pure. Now $j\colon \phi_\rho(A) \to \bigoplus_{\mu \sim \rho} \phi_\mu(A)$ is, by the definition of equivalence, a *-isomorphism, and $\tau \to \tau \circ j^{-1} = \tau'$ (transpose of τ) is a linear isometry of the corresponding dual spaces, under which $\omega_z \mid \phi_\rho(\mathfrak{A})$ is carried into $\omega_z \mid \bigoplus_{\mu \sim \rho} \phi_\mu(\mathfrak{A})$, which latter is therefore pure on its domain. Since, by [3], $(\bigoplus_{\mu \sim \rho} \phi_\mu(\mathfrak{A}))^-$ is a direct summand of $\Gamma(\mathfrak{A})^-$, with central projection P, say, we have $\omega_z(I - P) = 0$, and ω_z is a pure state of $\Gamma(\mathfrak{A})$. But $\omega_z(E) = 1$, so that $s(\omega_z) \subset E$, and $s(\omega_z)$ is minimal, concluding the proof.

We will call Γ *the atomic representation* of $\mathfrak{A}$. We define the central support of a state ρ to be $\inf\{P \mid P$ is a central projection in $\mathfrak{A}^-$ and $\rho(P) = 1\}$, or, equivalently, $\inf\{P \mid P$ is a central projection in $\mathfrak{A}^-$ and $s(\rho) \subset P\}$. Clearly, the central support of a pure state is a minimal central projection. The following is obvious:

PROPOSITION 2.5. *Two pure states are equivalent if and only if they have the same central support.*

Chapter 3.

Positive inner automorphisms. Let $\mathcal{H}$ be a Hilbert space, and T a positive operator, invertible in $\mathfrak{L}(\mathcal{H})$; then $\sigma(T)$ is contained in some interval, $[a, b]$, a and b both positive. Thus the real logarithm function restricts to a continuous function on $\sigma(T)$, defining by spectral theory the self-adjoint operator $\log T = L$ with the property $\exp(sL) = T^s$ for all real numbers s, and the mapping $s \to T^s$ is an analytic one-parameter group of positive definite operators on $\mathcal{H}$. We write "τ^s" for the inner-automorphism

$A \to T^{-s}AT^s$ of $\mathfrak{L}(\mathcal{H})$, "$\mathfrak{L}^2(\mathcal{H})$" for the Banach algebra $\mathfrak{L}(\mathfrak{L}(\mathcal{H}))$, "$e$" for the identity element of $\mathfrak{L}^2(\mathcal{H})$, and "$\operatorname{ad} L$" for the map $A \to AL - LA$, A in $\mathfrak{L}(\mathcal{H})$.

LEMMA 3.0. *The mapping $A, B \to A \otimes B$, where $A \otimes B(X) = AXB$ $(X \in \mathfrak{L}(\mathcal{H}))$, is a continuous bilinear mapping of $\mathfrak{L}(\mathcal{H}) \times \mathfrak{L}(\mathcal{H})$ into $\mathfrak{L}^2(\mathcal{H})$.*

Proof. The bilinearity is trivial. Continuity follows from continuity at $(0,0)$ and the relation $\| AXB \| \leqq \| A \| \, \| X \| \, \| B \|$.

LEMMA 3.1. $\tau^s = \exp(s \cdot \operatorname{ad} L)$.

Proof. For $A \in \mathfrak{L}(\mathcal{H})$, we have $s^{-1}(\tau^s - e)(A) = s^{-1}(T^{-s}AT^s - A)$ $= s^{-1}(T^{-s}AT^s - AT^s) + s^{-1}(AT^s - A) = A \cdot s^{-1}(T^s - I) + s^{-1}(T^{-s} - I)AT^s$. That is, $s^{-1}(\tau^s - e) = I \otimes s^{-1}(T^s - I) + s^{-1}(T^{-s} - I) \otimes T^s$. Taking account of the facts $\lim_{s \to 0} T^s = I$, $\lim_{s \to 0} s^{-1}(T^s - I) = L$, and applying Lemma 3.0, we have $\lim_{s \to 0} s^{-1}(\tau^s - e) = I \otimes L - L \otimes I = \operatorname{ad} L$, in the norm topology of $\mathfrak{L}^2(\mathcal{H})$. The analytic groups $s \to \tau^s$ and $s \to \exp(s \cdot \operatorname{ad} L)$ have the same infinitesimal generator $\operatorname{ad} L$, and so coincide: $\tau^s = \exp(s \cdot \operatorname{ad} L)$, as claimed. [See, for instance, 5, p. 283, Theorem 9.4.2.]

LEMMA 3.2. *For each real number s, τ^s has a positive real spectrum.*

Proof. Since each τ^s is an inner-automorphism of $\mathfrak{L}(\mathcal{H})$ defined by a positive operator, it suffices to prove the lemma for $s = 1$. Now T is approximable in norm by a sequence Q_m of operators of the form $\sum_{i=1}^{n} \lambda_i E_i$, where $\lambda_i \in \sigma(T)$, each E_i is a spectral projection of T, and $\sum E_i = I$; moreover, the Q_m can be so chosen that, simultaneously, the corresponding operators $R_m = \sum_{i=1}^{n} \lambda_i^{-1} E_i$ (same E_i!) converge to T^{-1}. Then by Lemma 3.0, $R_m \otimes Q_m$ converges to $\tau = T^{-1} \otimes T$ in $\mathfrak{L}^2(\mathcal{H})$. Now

$$R_m \otimes Q_m = \sum_{i,j} \lambda_i^{-1} \lambda_j E_i \otimes E_j,$$

whereas the $E_i \otimes E_j$ are easily seen to be idempotents in $\mathfrak{L}^2(\mathcal{H})$ satisfying $(E_i \otimes E_j)(E_k \otimes E_l) = 0$ if $(i,j) \neq (k,l)$, $\sum_{i,j} E_i \otimes E_j = e$. It is immediate that the spectrum of $R_m \otimes Q_m$ is the set $\{\lambda_i^{-1}\lambda_j\}_{i,j}$, so that for all n, $\sigma(R_n \otimes Q_n) \subset [m, M]$, where $m = \min\{x^{-1}y \mid x, y \in \sigma(T)\}$, $M = \max\{x^{-1}y \mid x, y \in \sigma(T)\} = m^{-1}$. Noting that each $R_m \otimes Q_m$ commutes with τ, we apply the continuity of spectra [10, Theorem (1.6.17)] to conclude that for each

neighborhood V of zero in the complex plane, $\sigma(\tau) \subset [m, M] + V$. Then $\sigma(\tau) \subset [m, M]$, proving the lemma.

PROPOSITION 3.3. *There is a sequence of polynomials p_n with real coefficients, satisfying* $\lim_{n\to\infty} \| p_n(\tau) - \operatorname{ad} L \| = 0$.

Proof. If a is the midpoint of $[m, M]$, the real logarithm function extends to a function analytic in the disc $\Delta = \{z \mid |z - a| < \frac{1}{2}M\}$; we call the extension "log." The Taylor series for log about the point a affords a sequence of polynominals p_n with real coefficients converging uniformly in Δ to log. A standard result in the functional calculus for Banach algebras [5, Theorem 5.2.5] then provides that $p_n(\tau)$ converges in $\mathfrak{L}^2(\mathfrak{H})$ to a logarithm Λ of τ. An application of the spectral mapping theorem [5, Theorem 5.3.1] shows that Λ has a real spectrum, hence that $\exp(s\Lambda)$ has real spectrum for all real s.

Now let $\mathfrak{D}$ be the closed subalgebra of $\mathfrak{L}^2(\mathfrak{H})$ generated by $\operatorname{ad} L$ and e. $\mathfrak{D}$ contains $\exp(\operatorname{ad} L) = \tau$, hence all $p_n(\tau)$, and, finally, Λ. Moreover, $\mathfrak{D}$ is commutative. We apply a theorem of E. R. Lorch [8, p. 421; 5, Theorem 5.5.5] to the effect that in a commutative Banach algebra $\mathfrak{D}$, a period of exp is of the form $2\pi i \sum_{\nu=1}^{k} n_\nu j_\nu$, with j_ν idempotent in $\mathfrak{D}$, n_ν integers; this yields

$$\Lambda = \operatorname{ad} L + 2\pi i \sum_{\nu=1}^{k} n_\nu j_\nu. \tag{1}$$

We may, of course, assume "orthogonality": $j_\nu j_\mu = 0$, if $\nu \neq \mu$. We assume also that $n_\nu \neq 0$, $1 \leqq \nu \leqq k$. Then for each s in R,

$$\exp(s\Lambda) = \exp(s \cdot \operatorname{ad} L) \cdot \left(e - \sum_{\nu=1}^{k} j_\nu + \sum_{\nu=1}^{k} e^{2\pi i s n_\nu} j_\nu\right). \quad (\log e = 1.) \tag{2}$$

Now, if $0 < |s| < \min_\nu\{1/(2|n_\nu|)\}$, $e^{2\pi i s n_\nu}$ is non-real. If h is an arbitrary homomorphism of $\mathfrak{D}$ onto the complex numbers, we note as a consequence of $h(e) = h(\sum j_\nu + (e - \sum j_\nu)) = 1$, that $h(e - \sum j_\nu) = 1$. (Recall that each of $h(\exp(s\Lambda))$, $h(\exp(s \cdot \operatorname{ad} L))$ is real, and in fact positive!)

Applying h to (2), we now have $h(\exp(s\Lambda)) = h(\exp(s \cdot \operatorname{ad} L))$; differentiating, we obtain $h(\Lambda) = h(\operatorname{ad} L)$ for every homomorphism h of $\mathfrak{D}$ onto the complexes. But then $\Lambda - \operatorname{ad} L \in \text{radical}(\mathfrak{D})$, while from (1), $\Lambda - \operatorname{ad} L = 2\pi i \sum_{\nu=1}^{k} n_\nu j_\nu$. Because the radical is an ideal, this implies each j_ν lies in radical $(\mathfrak{D})$; but an idempotent in the radical must be zero. We conclude that $\Lambda = \operatorname{ad} L$, and Proposition 3.3 is proved.

COROLLARY 3.4. *$\mathfrak{D}$ is generated, as a closed subalgebra of $\mathfrak{L}^2(\mathfrak{H})$, by τ.*

THEOREM 3.5 (*Invariance Theorem*). *Let $\mathcal{S}$ be a norm-closed linear subspace of $\mathfrak{L}(\mathcal{H})$, T a positive invertible operator in $\mathfrak{L}(\mathcal{H})$. Then if $\mathcal{S}$ is invariant under $T^{-1}\otimes T$, $\mathcal{S}$ is invariant also under* $A\to A\log T-(\log T)A$, *and under $T^{-s}\otimes T^{s}$ for all s in R.*

Proof. The set $\mathfrak{L}_{\mathcal{E}}(\mathcal{S})$ of bounded linear operators on a Banach space $\mathcal{E}$, leaving invariant a closed linear subspace $\mathcal{S}$ of $\mathcal{E}$, is a closed subalgebra of $\mathfrak{L}(\mathcal{E})$. Therefore, by Corollary 3.4, $\mathfrak{L}_{\mathfrak{L}(\mathcal{H})}(\mathcal{S})$ contains $\mathfrak{D}$.

Chapter 4.

Structure of isomorphisms.

PROPOSITION 4.1. *Let ψ be an isomorphism of C^*-algebra $\mathfrak{A}$ onto C^*-algebra $\mathfrak{B}$. Then*

i) *ψ is bounded;*

ii) *the double transpose ψ'' of ψ is an isomorphism of $\mathfrak{A}''$ onto $\mathfrak{B}''$.*

Proof of i). We follow the outline of [1, p. 15, ex. 5]. For $B\in\mathfrak{B}$, put $\|B\|_1=\|\psi^{-1}(B)\|$. We call $\|\cdot\|_1$ the "one-norm."

a) $\|B\|^2\leqq\|B^*\|_1\|B\|_1$. In fact, if $\lambda>\|B^*B\|_1$,

$$\|I-(I-\lambda^{-1}B^*B)\|_1=\lambda^{-1}\|B^*B\|_1<1,$$

and $\sum_{n=0}^{\infty}(\lambda^{-1}B^*B)^n$ converges absolutely in the one-norm to an inverse for $(I-\lambda^{-1}B^*B)$ in $\mathfrak{B}$. Thus $\sigma(B^*B)\subset\Delta_{\|B^*B\|_1}$, the disc of radius $\|B^*B\|_1$. But since B^*B is a positive operator, $\|B^*B\|=\max\{\lambda\mid\lambda\in\sigma(B^*B)\}$, so that $\|B^*B\|\leqq\|B^*B\|_1$. Finally, $\|B\|^2=\|B^*B\|\leqq\|B^*B\|_1\leqq\|B^*\|_1\|B\|_1$, or $\|B\|^2\leqq\|B^*\|_1\|B\|_1$, as claimed.

b) Next, we show that * is continuous in the one-norm on $\mathfrak{B}$. Since * is a one-one linear mapping of the real Banach space $(\mathfrak{B},\|\cdot\|_1)$ onto itself, it suffices by the closed-graph theorem to show that the graph of * is closed. Then let $\|X_n-Y\|_1\to0$, $\|X_n^*-Z\|_1\to0$ in $\mathfrak{B}$. We have, using the inequality of a) above,

$$\|X_n^*-Y^*\|^2=\|X_n-Y\|^2\leqq\|X_n^*-Y^*\|_1\|X_n-Y\|_1.$$

Since X_n^* is Cauchy in the one-norm, the first factor on the right is bounded, while the second converges to 0; thus $\|X_n^*-Y^*\|^2\to0$. On the other hand, $\|X_n^*-Z\|^2\leqq\|X_n-Z^*\|_1\|X_n^*-Z\|_1$, and by an alalogous argument, $\|X_n^*-Z\|^2\to0$. Thus $Z=Y^*$, proving the graph of * closed.

c) ψ is continuous provided $\| Y_n \|_1 \to 0$ implies $\| Y_n \| \to 0$ for a sequence (Y_n) in $\mathfrak{B}$. Using b) we have $\| Y_n \|_1 \to 0$ implies $\| Y_n^* \|_1 \to 0$, and $\| Y_n \|^2 \leqq \| Y_n^* \|_1 \| Y_n \|_1 \to 0$, completing the proof of i) : ψ is bounded.

Once we remark that the weak-* and weak-operator topologies coincide on $\mathfrak{A}''$ (resp. $\mathfrak{B}''$), it is immediate from classical properties of the transpose that ψ'' is continuous with respect to the weak-operator topologies on $\mathfrak{A}''$ and $\mathfrak{B}''$. We use this, together with the weak continuity of the maps $A \to AB$ (B fixed) and $B \to AB$ (A fixed), to argue by approximation. ($\mathfrak{A}$ is weakly dense in $\mathfrak{A}''$.)

(We drop the double prime, hereafter writing "ψ" alike for the weak-weak continuous linear mapping of $\mathfrak{A}''$ onto $\mathfrak{B}''$ or for its restriction to $\mathfrak{A}$, canonically embedded in $\mathfrak{A}''$.)

Proof of ii). First, fix $B \in \mathfrak{A}$, and let A_ν be a net in $\mathfrak{A}$ weakly converging to $A \in \mathfrak{A}''$. Then $\psi(A_\nu) \rightharpoonup \psi(A)$, $A_\nu B \rightharpoonup AB$, $\psi(A_\nu B) \rightharpoonup \psi(AB)$, while $\psi(A_\nu B) = \psi(A_\nu)\psi(B) \rightharpoonup \psi(A)\psi(B)$. Thus for $B \in \mathfrak{A}$, $A \in \mathfrak{A}''$, $\psi(AB) = \psi(A)\psi(B)$. Now apply this for fixed $A \in \mathfrak{A}''$, and B_ν in $\mathfrak{A}$, $B_\nu \rightharpoonup B$, B arbitrary in $\mathfrak{A}''$. Then again, $\psi(B_\nu) \rightharpoonup \psi(B)$, $AB_\nu \rightharpoonup AB$, $\psi(AB_\nu) \rightharpoonup \psi(AB)$, while $\psi(AB_\nu) = \psi(A)\psi(B_\nu) \rightharpoonup \psi(A)\psi(B)$, whence finally, $\psi(AB) = \psi(A)\psi(B)$ for A, B arbitrary in $\mathfrak{A}''$. Since only the multiplicative property was at stake, ii) is proved.

COROLLARY 4.2. *An isomorphism ψ is a *-isomorphism if and only if* $\| \psi \| = 1$ [9, Cor. 4.8.19].

Proof. We know that a *-isomorphism is of norm one. If now ψ is an isomorphism of norm one, so is its extension to $\mathfrak{A}''$. Then if E is a projection in $\mathfrak{A}''$, $\psi(E)$ is an idempotent of norm one: that is, a projection; furthermore, $EF = 0$ implies $\psi(E)\psi(F) = \psi(EF) = 0$, so that ψ preserves orthogonality of projections. For A self-adjoint in $\mathfrak{A}$, approximate A by sums $\sum \alpha_i E_i$, the E_i mutually orthogonal projections; α_i real; the images are self-adjoint, and converge to $\psi(A)$, which is therefore self-adjoint. Since ψ preserves self-adjointness and is linear, ψ is a *-map, as claimed.

Remark. Since the weak and ultraweak topologies coincide on $\mathfrak{A}''$, $\mathfrak{B}''$ the restriction of ψ to the center of $\mathfrak{A}''$ is a normal *-isomorphism onto the center of $\mathfrak{B}''$. (The fact that an isomorphism of commutative C^*-algebras is a *-isomorphism is immediate from the fact that the spectrum is a topological-algebraic invariant.)

It is clear that an isomorphism ψ of $\mathfrak{A}$ onto $\mathfrak{B}$ preserves maximal left ideals, and so serves to define, via the one-one correspondence $\mathfrak{J}_\rho \leftrightarrow \rho$ described in Chapter 2 between maximal left ideals and pure states, a one-one mapping of the pure states of $\mathfrak{A}$ onto those of $\mathfrak{B}$; namely, $\rho \to \rho'$ precisely if $\psi(\mathfrak{J}_\rho) = \mathfrak{J}_{\rho'}$.

PROPOSITION 4.3. *Under the mapping $\rho \to \rho'$, equivalence of pure states is preserved.*

Proof. Since two pure states are equivalent precisely if their central supports are equal, it is required to show—in view of the normalcy of ψ on the center C of $\mathfrak{A}''$—only that (with P a projection in C) $\rho(P) = 1$ implies $\rho'(\psi(P)) = 1$. That is, if $E = s(\rho)$, and $E' = s(\rho')$, $E \subset P$ implies $E' \subset \psi(P)$. Now $E \subset P$ means $EP = PE = E$, so that $\psi(E)\psi(P) = \psi(P)\psi(E) = \psi(E)$; thus $\psi(P) \supset r(\psi(E))$. The proof will be concluded when we establish that $r(\psi(E)) = E'$. This we do in a series of lemmas:

LEMMA 4.3.1. *Let R be any ring, x and y in R, and suppose $xy = x$, $yx = y$; then $Rx = Ry$.*

Proof. $x = xy \in Ry$ implies $Rx \subset Ry$; by symmetry $Ry \subset Rx$, so that $Rx = Ry$.

LEMMA 4.3.2. *Let $\mathfrak{R}$ be a von Neumann algebra, $F \in \mathfrak{R}$ idempotent, and G the support of F. Then $\mathfrak{R}F = \mathfrak{R}G$.*

Proof. By the lemma above, we need only prove that $FG = F$ and $GF = G$. The first is obvious. But $G = s(F)$ means $G = r(F^*)$, and F^* is also idempotent. Thinking spatially, we verify for x in G and x in $(I - G)$ that $F^*G = G$; applying $*$, we conclude that $GF = G$, as desired.

LEMMA 4.3.3. *In the notation of Proposition 4.3, $r(\psi(E)) = E'$.*

Proof. Since by Corollary 2.3, $\mathfrak{A}''(I - E) = \mathfrak{J}_\rho^-$, we have

$$\begin{aligned}\mathfrak{B}'' \cdot s(\psi(I - E)) &= \mathfrak{B}''\psi(I - E) = \psi(\mathfrak{A}''(I - E)) \\ &= \psi(\mathfrak{J}_\rho^-) = \psi(\mathfrak{J}_\rho)^- = \mathfrak{J}_{\rho'}^- = \mathfrak{B}''(I - E').\end{aligned}$$

(The first equation uses Lemma 4.3.2. The third follows from the weak-weak bicontinuity of ψ.) Now applying the uniqueness aspect of Theorem 2.2, we conclude that $I - E' = s(\psi(I - E))$, or

$$E' = I - s(\psi(I - E)) = n(\psi(I - E)) = n(I - \psi(E)) = r(\psi(E)),$$

proving the lemma, and thereby Proposition 4.3.

We note, as a consequence of the fact that the two idempotents generate the same right ideal,[3] the relations

$$E'\psi(E) = \psi(E), \tag{1}$$

$$\psi(E)E' = E'. \tag{2}$$

Now fix a pure state ρ of $\mathfrak{A}$, and denoting as above $\psi(\mathfrak{J}_\rho)$ by $\mathfrak{J}_{\rho'}$, consider the linear mapping S_ρ of $\mathfrak{A}/\mathfrak{J}_\rho$ onto $\mathfrak{B}/\mathfrak{J}_{\rho'}$ defined by $A + \mathfrak{J}_\rho \to \psi(A) + \mathfrak{J}_{\rho'}$.

LEMMA 4.4. $\| S_\rho \| \leqq \| \psi \|$.

Proof. Let $E = s(\rho)$. We observe that, for $A \geqq 0$ in $\mathfrak{A}''$, $\rho(A) = \| EAE \|$. In fact, $\rho(A) = \rho(EAE)$, and since E is a minimal projection, and $E\mathfrak{A}''E$, as a von Neumann algebra, is generated by its projections, $E\mathfrak{A}''E$ is one-dimensional, linearly spanned by E. Then $\rho\,|_{E\mathfrak{A}''E}$ is completely determined by $\rho(E) = 1 = \| E \|$, and each of ρ, $\| \cdot \|$ is positively homogeneous, which proves the observation. Now,

$$\begin{aligned}\| S_\rho(A + \mathfrak{J}_\rho) \|_{\rho'}{}^2 &= \| \psi(A) + \mathfrak{J}_{\rho'} \|_{\rho'}{}^2 \\ &= \rho'(\psi(A)^*\psi(A)) \\ &= \| E'\psi(A)^*\psi(A)E' \| \\ &= \| \psi(A)E' \|^2 \\ &= \| \psi(A)\psi(E)E' \|^2 \text{ (using (2) above)} \\ &= \| \psi(AE)E' \|^2 \leqq \| \psi(AE) \|^2 \leqq \| \psi \|^2 \| AE \|^2.\end{aligned}$$

Since, even more easily, we have $\| A + \mathfrak{J}_\rho \|_\rho{}^2 = \| AE \|^2$, we conclude that $\| S_\rho(A + \mathfrak{J}_\rho) \|_{\rho'} \leqq \| \psi \| \, \| A + \mathfrak{J}_\rho \|_\rho$, or $\| S_\rho \| \leqq \| \psi \|$, as claimed.

Let now $\mathcal{E} \subset \mathfrak{P}(\mathfrak{A})$ be "complete"—that is, let $\mathcal{E}$ contain a complete set of representatives of the equivalence classes of pure states of $\mathfrak{A}$. Then it follows from Prop. 4.3 that $\mathcal{E}' = \{\rho' \mid \rho \in \mathcal{E}\}$ is complete. We take direct sums:

$$\mathcal{H} = \bigoplus_{\rho \in \mathcal{E}} \mathfrak{A}/\mathfrak{J}_\rho, \qquad \mathcal{K} = \bigoplus_{\rho' \in \mathcal{E}'} \mathfrak{B}/\mathfrak{J}_{\rho'}, \qquad S = \bigoplus_{\rho \in \mathcal{E}} S_\rho : \mathcal{H} \to \mathcal{K},$$

$$\Delta = \bigoplus_{\rho \in \mathcal{E}} \phi_\rho : \mathfrak{A} \to \mathfrak{L}(\mathcal{H}), \qquad \Delta' = \bigoplus_{\rho' \in \mathcal{E}'} \phi_{\rho'} : \mathfrak{B} \to \mathfrak{L}(\mathcal{K}).$$

We identify $\mathfrak{A}$ (resp. $\mathfrak{B}$) with its $*$-isomorphic image under Δ (resp. Δ') in $\mathfrak{L}(\mathcal{H})$ (resp. $\mathfrak{L}(\mathcal{K})$).

PROPOSITION 4.5. $S \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. *For each* $A \in \mathfrak{A}$, $\psi(A) = SAS^{-1}$.

[3] $\mathfrak{B}''E' = \mathfrak{B}''\psi(E)^*$, so $(\mathfrak{B}''E')^* = (\mathfrak{B}''\psi(E)^*)^*$: i.e., $E'\mathfrak{B}'' = \psi(E)\mathfrak{B}''$.

Proof. $\| S \| = \sup_{\rho \in \mathcal{E}} \| S_\rho \| \leqq \| \psi \|$, by Lemma 4.4. The second statement is proved by straightforward computation:

$$\begin{aligned}\psi(A)(\psi(B) + \mathfrak{J}_{\rho'}) &= \psi(A)\psi(B) + \mathfrak{J}_{\rho'} = \psi(AB) + \mathfrak{J}_{\rho'} \\ &= S(AB + \mathfrak{J}_\rho) = SA(B + \mathfrak{J}_\rho) = SAS^{-1}(\psi(B) + \mathfrak{J}_{\rho'}).\end{aligned}$$

Extend ψ to $\mathfrak{L}(\mathcal{H})$ by $\psi = S \otimes S^{-1}$.

THEOREM A. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be C^*-algebras, $\mathcal{H}$ and $\mathcal{K}$ the atomic spaces of $\mathfrak{A}$, $\mathfrak{B}$ respectively. Then every isomorphism of $\mathfrak{A}$ onto $\mathfrak{B}$ can be extended to an isomorphism of $\mathfrak{L}(\mathcal{H})$ onto $\mathfrak{L}(\mathcal{K})$ of the form $A \to SAS^{-1}$ for some $S \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Eevery $*$-isomorphism can be so extended, with S unitary.*

Proof. With $\mathcal{E} = \mathfrak{P}(\mathfrak{A})$, the first statement is immediate from Prop. 4.5. The second statement follows from the easily verified fact that if ψ is a $*$-isomorphism, then $\rho' = \rho \circ \psi^{-1}$, so that each S_ρ, hence finally S itself, is unitary.

Now, retaining the notation above, put $T = S^*S \in \mathfrak{L}(\mathcal{H})$; then the polar decomposition of S gives $S = VT^{\frac{1}{2}}$ or $V = ST^{-(\frac{1}{2})} = (S^*)^{-1}T^{\frac{1}{2}}$, with $V : \mathcal{H} \to \mathcal{K}$ unitary.

LEMMA 4.6. $V\mathfrak{A}V^* = \mathfrak{B}$.

Proof. Since for $A \in \mathfrak{A}$, $VAV^* = ST^{-(\frac{1}{2})}AT^{\frac{1}{2}}S^{-1} = \psi(T^{-(\frac{1}{2})}AT^{\frac{1}{2}})$ by Prop. 4.5, it suffices to show that $T^{-(\frac{1}{2})}\mathfrak{A}T^{\frac{1}{2}} = \mathfrak{A}$. But the Invariance Theorem of Chapter 3 reduces this to showing $T^{-1}\mathfrak{A}T = \mathfrak{A}$. We have, with $A \in \mathfrak{A}$:

$$\begin{aligned}T^{-1}AT &= S^{-1}(S^*)^{-1}AS^*S = S^{-1}(SA^*S^{-1})^*S \\ &= S^{-1}(\psi(A^*)^*)S = \psi^{-1}(\psi(A^*)^*) \in \mathfrak{A},\end{aligned}$$

proving the lemma.

Recalling that $A \to VAV^*$ is a $*$-isomorphism, we have

THEOREM B. *Isomorphic C^*-algebras are $*$-isomorphic.*

Remarks. Application of Theorem A to the case $\mathfrak{A} = \mathfrak{B}$ gives the structure of all automorphisms of $\mathfrak{A}$; also that of the $*$-automorphisms.

Since a closed two-sided ideal in a C^*algebra is necessarily self-adjoint [11], the structure of continuous homomorphisms of one C^* algebra onto another is given by a trivial extension of Theorem A.

Added in proof. The proof of Proposition 3.3 can be simplified by observing that, since $\exp(s \cdot \operatorname{ad} L)$ has (positive) real spectrum for all s,

ad L has real spectrum. This follows easily from the spectral mapping theorem, and enables us to conclude from equation (1) that $\sum n_\nu j_\nu \in$ radical($\mathfrak{D}$). The proof is then concluded as before.

REFERENCES.

[1] J. Dixmier, *Les algèbres d'operateurs dans l'espace Hilbertien*, Paris, 1957.

[2] I. Gelfand and M. Neumark, "On the imbedding of normed rings into the ring of operators in Hilbert space," *Matematičeskii Sbornik Novaja Serija*, vol. 12 (1943), pp. 197-213.

[3] J. G. Glimm and R. V. Kadison, "Unitary operators in C^*-algebras," *Pacific Journal of Mathematics*, vol. 10 (1960), pp. 547-56.

[4] A. Grothendieck, "Un résultat sur le dual d'une C^*-algèbre," *Journal de Mathématiques Pures et Appliquées* (9), vol. 36 (1957), pp. 97-108.

[5] E. Hille and R. S. Phillips, *Functional analysis and semi-groups*, American Mathematical Society Colloquium Publication XXXI, Providence, 1957.

[6] R. V. Kadison, "Irreducible operator algebras," *Proceedings of the National Academy of Sciences of the U.S.A.*, vol. 43 (1957), pp. 273-76.

[7] ———, "Transformations of states in operator theory and dynamics," *to appear*.

[8] E. R. Lorch, "The theory of analytic functions in normed abelian vector rings," *Transactions of the American Mathematical Society*, vol. 54 (1943), pp. 414-25.

[9] C. E. Rickart, *General theory of Banach algebras*, New York, 1960.

[10] S. Sakai, *The theory of W^*-algebras* (mimeographed notes), Yale University, 1962.

[11] I. E. Segal, "Two-sided ideals in operator algebras," *Annals of Mathematics*, vol. 50 (1949), pp. 856-865.

[12] S. Sherman, "The second adjoint of a C^*-algebra," *Proceedings of the International Congress of Mathematicians*, vol. 1 (1950), p. 470.

[13] Z. Takeda, "Conjugate spaces of operator algebras," *Proceedings of the Japan Academy*, vol. 30 (1954), pp. 90-95.

FIELDS, OPTIONALITY AND MEASURABILITY.*

By K. L. CHUNG and J. L. DOOB.

1. **Basic notions.** Throughout this paper the following notation will be used:

$$T^0 = (0, \infty), \qquad T = (-\infty, +\infty), \qquad T^* = [-\infty, +\infty],$$
$$\forall t \in T^*: \qquad T_t = (-\infty, t].$$

The usual Borel field on T, trivially extended to T^* if need be, is denoted by $\mathcal{B}$. Its restriction to T_t is denoted by $\mathcal{B}_t$.

N is the set of all integers; R is the set of all rational numbers; the restrictions of N and R to T^0 are denoted by N^0 and R^0 respectively. Where it is not specified, the letters s, t, u denote elements of T^0 or T; the letters k, m, n elements of N^0 or N; depending on the context. The quantifier "$\forall t$" or "$\forall n$" will be omitted sometimes.

For two numbers or two numerical functions α, β with the same domain, we write

$$\alpha \wedge \beta = \min(\alpha, \beta), \qquad \alpha \vee \beta = \max(\alpha, \beta).$$

The lattice notation $\wedge$ and $\vee$ will also be used for Borel fields. In this case if $\{\mathcal{F}_i\}$ is any indexed family of Borel fields on the same set, we put

$$\bigwedge_i \mathcal{F}_i = \text{the largest Borel field contained in every } \mathcal{F}_i;$$
$$\bigvee_i \mathcal{F}_i = \text{the smallest Borel field containing every } \mathcal{F}_i.$$

If Ω is an abstract set (space), a *Borel field* (B. F.) on Ω is a collection of subsets of Ω which is closed under complementation in Ω and countable union (and intersection). If $\Delta \subset \Omega$, we denote by

$$\Delta \cap \mathcal{F}$$

the collection of sets of the form $\Delta \cap F$ with F ranging over $\mathcal{F}$; this is seen to be a B. F. on Δ if Δ is not empty, otherwise it consists of the empty set.

If S is a subset of T^*, a family of B. F.'s $\{\mathcal{F}_s\}$ on Ω, indexed by S, is called nondecreasing iff

$$\forall s < t: \qquad \mathcal{F}_s \subset \mathcal{F}_t. \tag{1}$$

Received June 8, 1964.

* This research is supported in part by the Office of Scientific Research of the United States Air Force.

Such a family can be trivially enlarged to one on the index set T^*, and we may suppose this to have been done in what follows. The following notation will be used:

$$\mathcal{F}_{t-} = \bigvee_{s<t} \mathcal{F}_s, \qquad \mathcal{F}_{t+} = \bigwedge_{s>t} \mathcal{F}_s.$$

Let α be a function on Ω to T^*. We shall write

$$\alpha \in \mathcal{F}$$

iff α is measurable with respect to the B. F. $\mathcal{F}$ in the usual sense, and say "α belongs to (or is contained in) $\mathcal{F}$" and "$\mathcal{F}$ contains α." This simplification of language seems overdue. The smallest B. F. containing a collection of functions such as $\{x_s, s \in S\}$ is also said to be *generated by* it (or them) and denoted by $\mathcal{F}\{x_s, s \in S\}$. In case of a single function, say α, the notation becomes $\mathcal{F}\{\alpha\}$. These definitions have their obvious generalizations if the range of the functions is in an abstract space, as will be supposed in §§ 3-4.

From now on, a B. F. is on Ω and a function is on Ω to T^*, unless otherwise specified. The set $\{\omega : \alpha(\omega) > t\}$, e. g., will be abbreviated as $\{\alpha > t\}$.

Definition 1. Let $\{\mathcal{F}_t, t \in T\}$ be a nondecreasing family of B. F.'s and α a function. The B. F. generated by the collections:

$$\{\alpha > t\} \cap \mathcal{F}_t, \qquad t \in T, \tag{2}$$

will be denoted by $\mathcal{F}_{\alpha-}$.

PROPOSITION 1. *An equivalent definition of $\mathcal{F}_{\alpha-}$ is obtained if we replace $\mathcal{F}_t$ in* (2) *by $\mathcal{F}_{t-}$ or $\mathcal{F}_{t+}$; or if we replace* (2) *by*

$$\{\alpha \geqq t\} \cap \mathcal{F}_{t-}, \qquad t \in T. \tag{2'}$$

Finally we may replace T in (2) *or* (2′) *by any dense subset of T.*

Proof. In this proof let us denote the B. F. obtained with $\mathcal{F}_{t-}$, $\mathcal{F}_t$, $\mathcal{F}_{t+}$ in (2) respectively by $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$; and the B. F. generated by the sets in (2′) by $\mathcal{B}_4$. It follows from (1) that

$$\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{B}_3.$$

On the other hand, we have for each $\Lambda \in \mathcal{F}_{t+}$,

$$\{\alpha > t\} \cap \Lambda = \bigcup_{n=1}^{\infty} [\{\alpha > t + 1/n\} \cap \Lambda]. \tag{3}$$

Clearly each member of the union above belongs to $\mathcal{B}_1$. Hence $\mathcal{B}_3 \subset \mathcal{B}_1$ and we have proved the equality of $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$.

Next, the equation (3) remains valid if the "$>$" on the right side is replaced by "$\geqq$," and now each member of the union there belongs to $\mathcal{B}_4$; hence $\mathcal{B}_3 \subset \mathcal{B}_4$. Conversely, for each m and $\Lambda \in \mathcal{F}_{t-(1/m)}$:

$$\{\alpha \geqq t\} \cap \Lambda = \bigcap_{n=m}^{\infty} [\{\alpha > t - (1/n)\} \cap \Lambda]$$

where each member of the intersection belongs to $\mathcal{B}_3$. Now it follows from a known result (see e.g. [5; p. 25]) that if $\Delta \subset \Omega$ and $\{\mathcal{F}_i\}$ is any collection of B.F.'s, we have

$$\bigvee_i (\Delta \cap \mathcal{F}_i) = \Delta \cap (\bigvee_i \mathcal{F}_i). \tag{4}$$

Hence $\forall t$:

$$\bigvee_{m=1}^{\infty} (\{\alpha \geqq t\} \cap \mathcal{F}_{t-(1/m)}) = \{\alpha \geqq t\} \cap \mathcal{F}_{t-}.$$

Consequently $\mathcal{B}_4 \subset \mathcal{B}_3$ and we have proved the equality of $\mathcal{B}_3$ and $\mathcal{B}_4$.

Finally, the last assertion of Proposition 1 follows from the equation: for each $\Lambda \in \mathcal{B}_t$,

$$\{\alpha > t\} \cap \Lambda = \bigcup_{t < r \in R} [\{\alpha > r\} \cap \Lambda]$$

if R is a dense subset of T.

PROPOSITION 2. *We have*

$$\mathcal{F}\{\alpha\} \subset \mathcal{F}_{\alpha-} \subset \mathcal{F}\{\alpha\} \vee \mathcal{F}_{+\infty}. \tag{5}$$

For each $B \in (t,\infty) \cap \mathcal{B}$, *we have*

$$\{\alpha \in B\} \cap \mathcal{F}_t \subset \mathcal{F}_{\alpha-}. \tag{6}$$

Proof. The first assertion is trivial but observe that we do not assume $\alpha \in \mathcal{F}_{+\infty}$. To prove (6) we note that

$$\forall u > t: \{\alpha > u\} \cap \mathcal{F}_t \subset \{\alpha > u\} \cap \mathcal{F}_u \subset \mathcal{F}_{\alpha-}.$$

Thus (6) is true if $B = (u,\infty)$ with $u > t$; hence it is true as asserted since the collection of sets for which it is true forms a Borel field on (t,∞).

PROPOSITION 3. *If* $\alpha \leqq \beta$, *then* $\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}$ *if and only if* $\alpha \in \mathcal{F}_{\beta-}$. *If* $\alpha_n \uparrow \alpha$[1] *and* $\forall n$: $\alpha_n \in \mathcal{F}_{\alpha-}$, *then*

$$\bigvee_n \mathcal{F}_{\alpha_n-} = \mathcal{F}_{\alpha-} \tag{7}$$

---

[1] The symbols $\uparrow$ and $\downarrow$ are used for monotone convergence in the non-strict sense.

Proof. We have if $\alpha \leqq \beta$:

$$\forall t: \{\alpha > t\} \cap \mathcal{F}_t = \{\alpha > t\} \cap [\{\beta > t\} \cap \mathcal{F}_t]. \tag{8}$$

If $\alpha \in \mathcal{F}_{\beta-}$, then $\{\alpha > t\} \in \mathcal{F}_{\beta-}$ and the set on the right side above belongs to $\mathcal{F}_{\beta-}$. This proves $\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}$. Conversely if the last inclusion holds, then by (5) : $\alpha \in \mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}$. The first assertion of Proposition 3 is proved. It follows that we have " $\subset$ " instead of " $=$ " in (7). On the other hand, we have for each t and $\Lambda \in \mathcal{F}_t$:

$$\{\alpha > t\} \cap \Lambda = \bigcup_n [\{\alpha_n > t\} \cap \Lambda] \subset \bigvee_n \mathcal{F}_{\alpha_n-}.$$

Hence (7) is proved.

PROPOSITION 3.1. *If $\alpha_n \uparrow \infty$ and if $\forall n: \alpha_n \in \mathcal{F}_{+\infty}$, then*

$$\bigvee_n \mathcal{F}_{\alpha_n-} = \mathcal{F}_{+\infty}.$$

Definition 2. We define

$$\mathcal{F}_{\alpha+} = \bigwedge_{\delta > 0} \mathcal{F}_{(\alpha+\delta)-}, \tag{9}$$

where the convention $\pm\infty + \delta = \pm\infty$ is used. It is obvious, even without the use of Proposition 3, that $\mathcal{F}_{(\alpha+\delta)-}$ is monotone in δ. Clearly $\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\alpha+}$.

Remark. If we define $\mathcal{F}_{\alpha^*}$ to be the B. F. generated by the collections

$$\{\alpha \geqq t\} \cap \mathcal{F}_t, \qquad t \in T,$$

it is easy to see that

$$\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\alpha^*} \subset \mathcal{F}_{\alpha+} = \bigwedge_{\delta>0} \mathcal{F}_{(\alpha+\delta)^*}.$$

We shall not use the intermediate B. F. $\mathcal{F}_{\alpha^*}$ except to note that if α is a constant t_0, then $\mathcal{F}_{\alpha-}$, $\mathcal{F}_{\alpha^*}$, $\mathcal{F}_{\alpha+}$ reduce respectively to $\mathcal{F}_{t_0-}$, $\mathcal{F}_{t_0}$, $\mathcal{F}_{t_0+}$. There are simple examples in which these three B. F.'s are in strictly increasing order.

Up to now the function α is arbitrary. We shall now turn our attention to a specially interesting class of functions.

Definition 3. The function α is called *optional relative to the non-decreasing family* $\{\mathcal{F}_t, t \in T\}$ of B. F.'s iff

$$\forall t \in T: \{\alpha < t\} \in \mathcal{F}_t; \tag{10}$$

it is called *strictly optional* iff

$$\forall t \in T: \{\alpha \leqq t\} \in \mathcal{F}_t. \tag{11}$$

If $\alpha \geqq 0$, then the index set T in the above may be replaced by T^0.

Proposition 5. *An equivalent definition of optionality is obtained if we replace* $\mathcal{F}_t$ *in* (10) *by* $\mathcal{F}_{t-}$ *or* $\mathcal{F}_{t+}$; *or if we replace* $\mathcal{F}_t$ *in* (11) *by* $\mathcal{F}_{t+}$; *or if we replace* T *by any dense subset of* T.

The proof is similar to that of Proposition 1.

It follows from the definition that $\alpha \in \mathcal{F}_{+\infty}$, and that

$$\forall t \in T:\ \alpha^{-1}(\mathcal{B}_t) \subset \mathcal{F}_{t+} \text{ or } \mathcal{F}_t \tag{12}$$

according as α is optional or strictly optional. The next proposition is also trivial.

Proposition 6. *Strict optionality implies optionality and the two notions coincide for a given* $\{\mathcal{F}_t\}$ *if and only if*

$$\forall t \in T:\ \mathcal{F}_t = \mathcal{F}_{t+}. \tag{13}$$

Furthermore, optionality relative to $\{\mathcal{F}_t\}$ *is equivalent to strict optionality relative to* $\{\mathcal{F}_{t+}\}$.

Unless otherwise specified, we shall regard the nondecreasing family $\{\mathcal{F}_t\}$ as given and optionality as relative to it. Clearly a constant function is strictly optional. The next two propositions, as well as Proposition 15 later, give ways of deriving new optional functions from given ones.

Proposition 7. *If* α *and* β *are both (strictly) optional, then so is* $\alpha \wedge \beta$ *and* $\alpha \vee \beta$; *similarly for a finite number of terms. If* $\forall n: \alpha_n$ *is optional, then so is each one of the following:*

$$\sup_n \alpha_n, \quad \inf_n \alpha_n, \quad \limsup_n \alpha_n, \quad \liminf_n \alpha_n. \tag{14}$$

If $\forall n: \alpha_n$ *is strictly optional, then so is the first one in* (14), *the others being optional but not necessarily strictly so.*

Proof. It is convenient to use one of the equivalent conditions given in Proposition 5. For instance, since

$$\{\sup_n \alpha_n \leqq t\} = \bigcap_n \{\alpha_n \leqq t\}, \qquad \{\inf_n \alpha_n < t\} = \bigcup_n \{\alpha_n < t\},$$

the assertions follow quickly by applying the appropriate criteria. As regards the last assertion of the proposition, we may take any α which is optional but not strictly so (Example 1 in § 5), and $\alpha_n = \alpha + 1/n$. Then α_n is strictly optional and

$$\alpha = \inf_n \alpha_n = \lim_n \alpha_n.$$

PROPOSITION 8. *If α and β are both optional and nonnegative, then $\alpha+\beta$ is optional. If furthermore one of the following three conditions is satisfied, then $\alpha+\beta$ is strictly optional:*

(i) $\alpha>0, \beta>0$;

(ii) $\beta>0$, *β is strictly optional*;

(iii) *α and β are both strictly optional.*

Remark. The condition "$\alpha>0$ and β is strictly optional" is not sufficient. Take α to be positive, optional but not strictly so; and take $\beta=0$.

Proof. Let α and β be optional and nonnegative throughout this proof. For each $t>0$:

$$\{\alpha+\beta<t\}=\bigcup_{r\in R\cap(0,t)}\{\alpha<r;\beta<t-r\}\in\mathfrak{F}_t,$$

proving the first assertion. Next, consider the decomposition for $t\geqq 0$:

$$\begin{aligned}&\{\alpha+\beta>t\}\\ &\quad=\{0<\alpha<t;\alpha+\beta>t\}\cup\{\alpha=0;\beta>t\}\\ &\qquad\cup\{\alpha>t;\beta=0\}\cup\{\alpha\geqq t;\beta>0\}.\end{aligned}$$

Let the four sets on the right side be denoted by Λ_1, Λ_2, Λ_3 and Λ_4 respectively. It is easy to see that for each $t>0$:

$$\Lambda_1=\bigcup_{r\in R\cap(0,t)}\{r<\alpha<t;\beta>t-r\}.$$

For each r in $(0,t)$ the set in $\{\ \}$ above belongs to $\mathfrak{F}_t$, hence $\Lambda_1\in\mathfrak{F}_t$.

Under (i), we have $\Lambda_2=0$, $\Lambda_3=0$, and if $t>0$:

$$\Lambda_4=\{\alpha\geqq t\}=\Omega-\{\alpha<t\}\in\mathfrak{F}_t.$$

Under (ii), we have $\Lambda_3=0$, and $\Lambda_4\in\mathfrak{F}_t$ as before. Furthermore if $t>0$:

$$\Lambda_2=\{\alpha\geqq 0\}\cap\{\beta>t\}\in\mathfrak{F}_{0+}\vee\mathfrak{F}_t=\mathfrak{F}_t.$$

Under (iii), we have if $t\geqq 0$:

$$\Lambda_2\in\mathfrak{F}_0\vee\mathfrak{F}_t=\mathfrak{F}_t,$$

and by symmetry $\Lambda_3\in\mathfrak{F}_t$. Furthermore

$$\Lambda_4=\{\alpha\geqq t\}\cap\{\beta>0\}\in\mathfrak{F}_t\vee\mathfrak{F}_0=\mathfrak{F}_t.$$

Hence, in cases (i) and (ii) we have

$$\{\alpha+\beta>t\}\in\mathfrak{F}_t$$

for every $t > 0$, while in case (iii) the same is true for every $t \geqq 0$. Proposition 8 is completely proved.

It should be observed that Proposition 8 is the only place where the nonnegativity of an optional function is supposed. In practice only the following trivial special case is needed.

PROPOSITION 8.1. *If α is optional and t is a positive constant, then $\alpha + t$ is strictly optional.*

PROPOSITION 9. *If α is optional, β is arbitrary and $\alpha \leqq \beta$, then*

$$\mathfrak{F}_{\alpha-} \subset \mathfrak{F}_{\beta-}, \qquad \mathfrak{F}_{\alpha+} \subset \mathfrak{F}_{\beta+}. \tag{15}$$

Proof. $\forall t$: $\{\alpha > t\} \in \mathfrak{F}_{t+}$, hence (8) with $\mathfrak{F}_t$ replaced by $\mathfrak{F}_{t+}$ proves the first relation in (15). Now $\forall \delta > 0$; $\alpha + \delta$ is (strictly) optional, hence

$$\mathfrak{F}_{(\alpha+\delta)-} \subset \mathfrak{F}_{(\beta+\delta)-}.$$

The second relation in (15) follows from this and Definition 2.

Definition 4. For any function α, let $\mathfrak{F}_{\alpha(+)}$ $[\mathfrak{F}_\alpha]$ denote the collection of subsets Λ of Ω for which $\Lambda \cap \{\alpha = +\infty\} \in \mathfrak{F}_{+\infty}$ and (16) [(17)] is true:

$$\forall t \in T: \Lambda \cap \{\alpha < t\} \in \mathfrak{F}_t; \tag{16}$$

$$\forall t \in T: \Lambda \cap \{\alpha \leqq t\} \in \mathfrak{F}_t. \tag{17}$$

Clearly $\mathfrak{F}_\alpha \subset \mathfrak{F}_{\alpha(+)} \subset \mathfrak{F}_{+\infty}$.

PROPOSITION 10. *If α is optional, then $\mathfrak{F}_{\alpha(+)}$ is a B.F.; if α is strictly optional, then also is $\mathfrak{F}_\alpha$. An equivalent definition of $\mathfrak{F}_{\alpha(+)}$ is obtained if we replace $\mathfrak{F}_t$ in (16) by $\mathfrak{F}_{t-}$ or $\mathfrak{F}_{t+}$; or if we replace $\mathfrak{F}_t$ in (17) by $\mathfrak{F}_{t+}$; or if we replace T by any dense subset of T.*

Proof. The first sentence follows from (10) and (11) by taking $\Lambda = \Omega$. The rest is similar to Proposition 1.

The following special case of the B.F.'s is instructive; the simple proof will be omitted.

PROPOSITION 11. *If α has a countable range $C \subset T$, then α is [strictly] optional if and only if*

$$\forall c \in C: \{\alpha = c\} \in \mathfrak{F}_{c+}[\mathfrak{F}_c];$$

and $\mathfrak{F}_{\alpha(+)}[\mathfrak{F}_\alpha, \mathfrak{F}_{\alpha-}]$ is the collection of all sets Λ of the form:

$$\Lambda = \bigcup_{c \in C} \{\alpha = c\} \cap \Lambda_c$$

where $\Lambda_c \in \mathfrak{F}_{c+}[\mathfrak{F}_c, \mathfrak{F}_{c-}]$.

PROPOSITION 12. *If* $\forall n$: α_n *is optional, and* $\alpha_n \downarrow \alpha$, *then*

$$\mathcal{F}_{\alpha(+)} = \bigwedge_n \mathcal{F}_{\alpha_n(+)}. \tag{18}$$

Remark. This is the counterpart of (7); note that (7) is valid if $\alpha_n \uparrow \alpha$ and each α_n is optional, by (5) and (15).

Proof. If α and β are optional, and $\alpha \leqq \beta$, $\Lambda \in \mathcal{F}_{\alpha(+)}$, then

$$\Lambda \cap \{\beta < t\} = [\Lambda \cap \{\alpha < t\}] \cap \{\beta < t\} \in \mathcal{F}_t,$$

by (16) and (10) for β. Hence $\mathcal{F}_{\alpha(+)} \subset \mathcal{F}_{\beta(+)}$. Applying this result to α and α_n, since α is optional by Proposition 7, we obtain " $\subset$ " instead of " $=$ " in (18). On the other hand, if Λ belongs to the right member of (18), then

$$\Lambda \cap \{\alpha < t\} = \bigcup_n [\Lambda \cap \{\alpha_n < t\}] \in \mathcal{F}_t.$$

Hence $\Lambda \in \mathcal{F}_{\alpha(+)}$ and (18) is proved.

PROPOSITION 13. *If* α *is arbitrary,* β *is [strictly] optional, and* $\alpha \leqq \beta$, *then*

$$\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta(+)} [\mathcal{F}_\beta]. \tag{19}$$

If α *is optional,* β *is arbitrary, and* $\alpha < \beta$,[2] *then*

$$\mathcal{F}_{\alpha(+)} \subset \mathcal{F}_{\beta-}. \tag{20}$$

Proof. The first assertion is proved by the formula:

$$[\mathcal{F}_t \cap \{t < \alpha\}] \cap \{\beta < u\} = [\mathcal{F}_t \cap \{t < \alpha < u\}] \cap \{\beta < u\},$$

which is a collection of sets contained in $\mathcal{F}_u$, by (6) if $t < u$, and trivially if $t \geqq u$. Hence each generating set of $\mathcal{F}_{\alpha-}$ belongs to $\mathcal{F}_{\beta(+)}$ by definition, proving (19); similarly for the strict case. To prove (20), let $\Lambda \in \mathcal{F}_{\alpha+}$ and put

$$\Lambda_r = \Lambda \cap \{\alpha < r\}. \tag{21}$$

Then $\Lambda_r \in \mathcal{F}_r$ and consequently

$$\Lambda = \bigcup_{r \in R} [\Lambda \cap \{\alpha < r < \beta\}] = \bigcup_{r \in R} [\Lambda_r \cap \{r < \beta\}] \in \mathcal{F}_{\beta-}.$$

PROPOSITION 14. *If* α *is optional, then*

$$\mathcal{F}_{\alpha(+)} = \mathcal{F}_{\alpha+}.$$

---

[2] This means $\alpha < \beta$ on $\{\alpha < +\infty\}$ and $\alpha = \beta$ on $\{\alpha = +\infty\}$.

Proof. We have by Proposition 12 and the first part of the Proposition 13:

$$\mathfrak{F}_{\alpha(+)} = \bigwedge_{n=1}^{\infty} \mathfrak{F}_{(\alpha+n^{-1})(+)} \supset \bigwedge_{n=1}^{\infty} \mathfrak{F}_{(\alpha+n^{-1})-}.$$

On the other hand, by the second part of Proposition 13:

$$\mathfrak{F}_{\alpha(+)} \subset \bigwedge_{n=1}^{\infty} \mathfrak{F}_{(\alpha+n^{-1})-}.$$

Together we obtain

$$\mathfrak{F}_{\alpha(+)} = \bigwedge_{n=1}^{\infty} \mathfrak{F}_{(\alpha+n^{-1})-} = \mathfrak{F}_{\alpha+}.$$

Proposition 14 is useful since one or the other definition proves more convenient in application. From now on we shall drop the notation $\mathfrak{F}_{\alpha(+)}$ in favor of $\mathfrak{F}_{\alpha+}$ (which is defined for every α). Let us remark that for a strictly optional α, we have $\mathfrak{F}_{\alpha\bullet} \subset \mathfrak{F}_{\alpha}$ but $\mathfrak{F}_{\alpha}$ need not coincide with $\mathfrak{F}_{\alpha\bullet}$ (Example 2 in § 5); we have $\mathfrak{F}_{\alpha} \subset \mathfrak{F}_{\beta}$ if β is also strictly optional and $\alpha \leqq \beta$.

PROPOSITION 15. *If α is optional, $\beta \in \mathfrak{F}_{\alpha+}$ and $\beta \geqq \alpha$, then β is optional. If furthermore either $\beta > \alpha$,*[3] *or α is strictly optional and $\beta \in \mathfrak{F}_{\alpha}$, then β is strictly optional.*

Proof. We prove the first strict version. $\forall t$: $\{\beta \leqq t\} \in \mathfrak{F}_{\alpha+}$, hence

$$\{\beta \leqq t\} = \{\beta \leqq t\} \cap \{\alpha < t\} \in \mathfrak{F}_t$$

by Definition 4. Hence β is strictly optional.

PROPOSITION 15.1. *If α is [strictly] optional, $\Delta \in \mathfrak{F}_{\alpha+}[\mathfrak{F}_{\alpha}]$, and*

$$\alpha_\Delta = \begin{cases} \alpha & \text{on } \Delta, \\ +\infty & \text{on } \Omega - \Delta; \end{cases}$$

then α_Δ is [strictly] optional.

For an application of the preceding proposition, see e.g. [2; Lemma, p. 34].

An important special case of Proposition 15 is as follows. For an arbitrary optional α, we define

$$\alpha_n = \frac{[2^n\alpha + 1]}{2^n}, \qquad n \in N^0. \tag{22}$$

Let Z be the set $\{m\, 2^{-n}\}$ where n ranges over N^0 and m over N. Then $\forall n$: α_n

[3] See preceding footnote.

has a countable range contained in Z and is strictly optional. Hence by Proposition 11, $\mathfrak{F}_{\alpha_n}$ consists of sets of the form

$$\bigcup_m [\{\alpha_n = m\,2^{-n}\} \cap \Lambda_{m2^{-n}}] = \bigcup_m [\{(m-1)2^{-n} \leqq \alpha < m\,2^{-n}\} \cap \Lambda_{m2^{-n}}]$$

where we use the notation (21). By the choice of the sequence $\{2^n\}$, we have $\alpha_{n+1} \leqq \alpha_n$ so that $\mathfrak{F}_{\alpha_{n+1}} \subset \mathfrak{F}_{\alpha_n}$, and $\mathfrak{F}_{\alpha+}$ is the B. F. intersection of the nonincreasing sequence $\{\mathfrak{F}_{\alpha_n}\}$, by (18). This type of approximation is useful in the evaluation of probabilities; see e.g. [1; pp. 165, 169].

2. Lattice and addition properties.

PROPOSITION 16. *If α is optional and β arbitrary, then*

$$(23) \qquad \{\alpha < \beta\} \cap \mathfrak{F}_{\alpha+} \subset \mathfrak{F}_{\beta-}, \qquad \{\alpha \leqq \beta\} \cap \mathfrak{F}_{\alpha+} \subset \mathfrak{F}_{\beta+}.$$

Proof. Let $\Lambda \in \mathfrak{F}_{\alpha+}$ and use the notation (21). We have

$$\Lambda \cap \{\alpha < \beta\} = \bigcup_{r \in R} [\Lambda \cap \{\alpha < r < \beta\}] = \bigcup_{r \in R} [\Lambda_r \cap \{r < \beta\}] \in \mathfrak{F}_{\beta-}$$

Next, $\forall m$:

$$\Lambda \cap \{\alpha \leqq \beta\} = \bigcap_{n=m}^{\infty} [\Lambda \cap \{\alpha < \beta + 1/n\}] \in \mathfrak{F}_{(\beta+m^{-1})-};$$

hence $\Lambda \cap \{\alpha \leqq \beta\}$ belongs to $\mathfrak{F}_{\beta+}$ by Definition 2.

PROPOSITION 17. *If α is optional and β arbitrary, then both $\{\alpha < \beta\}$ and $\{\alpha \leqq \beta\}$ belong to $\mathfrak{F}_{(\alpha\wedge\beta)+}$.*

Proof. Since

$$\{\alpha \leqq \beta\} = \{\alpha \leqq (\alpha \wedge \beta)\},$$

this set belongs to $\mathfrak{F}_{(\alpha\wedge\beta)+}$ by the second relation in (23). Applying this result to α and $\beta - n^{-1}$, we obtain

$$\{\alpha < \beta\} = \bigcup_n \{\alpha \leqq \beta - 1/n\} \in \bigvee_n \mathfrak{F}_{(\alpha\wedge(\beta-n^{-1}))+} \subset \mathfrak{F}_{(\alpha\wedge\beta)+}.$$

PROPOSITION 18. *If α and β are optional, then*

$$(24\pm) \qquad \{\alpha \leqq \beta\} \cap \mathfrak{F}_{(\alpha\wedge\beta)\pm} = \{\alpha \leqq \beta\} \cap \mathfrak{F}_{\alpha\pm};$$

$$(25\pm) \qquad \{\alpha \leqq \beta\} \cap \mathfrak{F}_{(\alpha\vee\beta)\pm} = \{\alpha \leqq \beta\} \cap \mathfrak{F}_{\beta\pm},$$

where in each formula we take "$+$" or "$-$" together, and similar relations also hold if "$\leqq$" is replaced by "$<$" everywhere.

Proof. It follows from (15) that we have " $\subset$ " in (24) and " $\supset$ " in (25). To prove " $\supset$ " in (24—), we need only observe that

$$\{\alpha \leqq \beta\} \cap [\mathcal{F}_t \cap \{t < \alpha\}] = \{\alpha \leqq \beta\} \cap [\mathcal{F}_t \cap \{t < (\alpha \wedge \beta)\}];$$

similarly for (25—), and with " $<$ " in place of " $\leqq$." To prove (24+) and (25+), we observe that, e.g.,

$$\{\alpha \leqq \beta\} \cap \mathcal{F}_{(\alpha \wedge \beta)+} = \{\alpha + \delta \leqq \beta + \delta\} \cap \bigwedge_{\delta > 0} \mathcal{F}_{[(\alpha \wedge \beta)+\delta]-}.$$

Applying (24—) and (25—) to $a + \delta$ and $\beta + \delta$, we obtain (24+) and (25+) by the following general result, which is a counterpart to (4). If $\Delta \subset \Omega$, and $\{\mathcal{F}_i\}$ is a nonincreasing sequence of B. F.'s, then

$$\bigwedge_i (\Delta \mathcal{F}_i) = \Delta \cap (\bigwedge_i \mathcal{F}_i). \tag{26}$$

Proposition 18 is proved.

If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two collections of subsets of Ω, we denote by $\mathcal{B}_1 \cup \mathcal{B}_2$ the collection of all sets of the form $\Lambda_1 \cup \Lambda_2$ where $\Lambda_1 \in \mathcal{B}_1$, $\Lambda_2 \in \mathcal{B}_2$.

PROPOSITION 19. *If α and β are optional, then*

$$\mathcal{F}_{(\alpha \vee \beta)+} = \mathcal{F}_{\alpha+} \cup \mathcal{F}_{\beta+}, \tag{27}$$

$$\mathcal{F}_{(\alpha \wedge \beta)+} = \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{\beta+}, \tag{28}$$

$$\mathcal{F}_{(\alpha \vee \beta)-} = \mathcal{F}_{\alpha-} \vee \mathcal{F}_{\beta-}, \tag{29}$$

but in general

$$\mathcal{F}_{(\alpha \wedge \beta)-} \neq \mathcal{F}_{\alpha-} \wedge \mathcal{F}_{\beta-}. \tag{30}$$

Remark. These relations extend at once to a finite number of optional functions by induction. Previous relations (7) and (18) are the limiting cases of (29) and (28) respectively. The limiting case of (27) is in general false, see Example 7 in § 5.

Proof. It follows from (15) that we have " $\supset$ " in (27) and (29), " $\subset$ " in (28) and (30). It remains to prove the opposite inclusion in the first three relations and disprove it in the fourth.

Let $\Lambda \in \mathcal{F}_{(\alpha \vee \beta)+}$, then by (25+) and Proposition 16:

$$\{\alpha \leqq \beta\} \cap \Lambda \in \{\alpha \leqq \beta\} \cap \mathcal{F}_{\beta+} \subset \mathcal{F}_{\beta+}.$$

Interchanging α and β in the above and taking the union of the two results we obtain (27) with " $\subset$."

Next, let $\Lambda \in \mathcal{F}_{\alpha+}$, then by (24+) and Proposition 17:

$$\{\alpha \leqq \beta\} \cap \Lambda \in \{\alpha \leqq \beta\} \cap \mathcal{F}_{(\alpha\wedge\beta)+} \subset \mathcal{F}_{(\alpha\wedge\beta)+}.$$

If also $\Lambda \in \mathcal{F}_{\beta+}$, then we may interchange α and β in the above and take the union. Hence each Λ belonging to the right member of (28) belongs also to the left member.

Next, we observe that by (5):

$$\{\alpha \leqq \beta\} \in \mathcal{F}\{\alpha\} \vee \mathcal{F}\{\beta\} \subset \mathcal{F}_{\alpha-} \vee \mathcal{F}_{\beta-}$$

Hence if $\Lambda \in \mathcal{F}_{(\alpha\vee\beta)-}$, then by (25—):

$$\{\alpha \leqq \beta\} \cap \Lambda \in \{\alpha \leqq \beta\} \cap \mathcal{F}_{\beta-} \subset \mathcal{F}_{\alpha-} \vee \mathcal{F}_{\beta-}.$$

Interchanging α and β we conclude as before that Λ belongs to the right member of (29).

Finally, an example of (30) will be given in Example 3 of § 5.

If α is optional, then so is $\alpha \wedge t$ for each $t \in T$, and $\alpha + t$ is strictly optional for each $t \in T^0$. Given $\{\mathcal{F}_t, t \in T\}$ and α, let us write:

$$\mathcal{E}_t = \mathcal{F}_{(\alpha\wedge t)+}, \qquad \mathcal{B}_t = \mathcal{F}_{\alpha+t}. \tag{31}$$

The two families $\{\mathcal{E}_t, t \in T\}$ and $\{\mathcal{B}_t, t \in T^0\}$ are both nondecreasing.

PROPOSITION 20. *If α is optional, then*

$$\bigvee_t \mathcal{F}_{(\alpha\wedge t)\pm} = \mathcal{F}_{\alpha\pm}. \tag{32$\pm$}$$

Proof. Since $\lim_{t\uparrow\infty} (\alpha \wedge t) = \alpha$, (32—) is just a special case of (7). However, the analogue of (7) with "—" replaced by "+" is in general false; see Example 7 in § 5.

To prove (32+), let $\Lambda \in \mathcal{F}_{\alpha+}$. Using the notation (21), we have for every $t \in T$:

$$\Lambda_n \cap \{(\alpha \wedge n) < t\} = \Lambda \cap \{\alpha < (n \wedge t)\} \in \mathcal{F}_{n\wedge t} \subset \mathcal{F}_t;$$

hence for every $n \in N^0$:

$$\Lambda_n \in \mathcal{F}_{(\alpha\wedge n)+}. \tag{33}$$

Next, we have

$$\{\alpha = +\infty\} \cap \mathcal{F}_{+\infty} \subset \mathcal{F}_{\alpha-}. \tag{34}$$

For if $M \in \mathcal{F}_n$ then $\{\alpha > n\} \cap M \in \mathcal{F}_{\alpha-}$ and so

$$\{\alpha = +\infty\} \cap M = \bigcap_n [\{\alpha > n\} \cap M] \in \mathcal{F}_{\alpha-}. \tag{35}$$

Since $\alpha \in \mathcal{F}_{\alpha-}$, the class of sets M for which (35) holds forms a B. F. This B. F. contains every $\mathcal{F}_n$ as just shown, hence it contains $\mathcal{F}_{+\infty}$, proving (34). Since also $\Lambda \in \mathcal{F}_{+\infty}$, it follows that $\{\alpha = +\infty\} \cap \Lambda$ belongs to $\mathcal{F}_{\alpha-}$, and consequently to the left member of (32—), by (32—), and *a fortiori* to that of (32+). Now it follows from this, (33) and the equation

$$\Lambda = [\{\alpha = +\infty\} \cap \Lambda] \cup (\bigcup_{n=1}^{\infty} \Lambda_n),$$

that Λ belongs to the left member of (32+). Thus we have "$\supset$" in (32+). The opposite inclusion follows from (15) and so (32+) is proved.

PROPOSITION 21. *If β as well as α is optional relative to $\{\mathcal{F}_t\}$, then β is optional relative to $\{\mathcal{E}_t\}$ if and only if $\beta \in \mathcal{F}_{\alpha+}$. This is the case if $\beta \leqq \alpha$.*

Proof. If β is optional relative to $\{\mathcal{E}_t\}$, then by (28):

$$\beta \in \mathcal{E}_t = \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} \subset \mathcal{F}_{\alpha+}.$$

Conversely if $\beta \in \mathcal{F}_{\alpha+}$, then $\{\beta < t\} \in \mathcal{F}_{t+}$, and so

$$\{\beta < t\} \in \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} = \mathcal{F}_{(\alpha \wedge t)+}.$$

If $\beta \leqq \alpha$, then $\beta \in \mathcal{F}_{\beta+} \subset \mathcal{F}_{\alpha+}$ by (15). Proposition 21 is proved.

PROPOSITION 22. *If β is optional relative to $\{\mathcal{E}_t\}$, then*

(36) $$\mathcal{E}_{\beta+} = \mathcal{F}_{(\alpha \wedge \beta)+}.$$

Proof. Let $\Lambda \in \mathcal{F}_{(\alpha\wedge\beta)+}$, then for every t:

$$\Lambda \cap \{\beta < t\} = [\Lambda \cap \{\alpha \wedge \beta) < t\}] \cap \{\beta < t\} \in \mathcal{F}_t \cap \{\beta < t\} \subset \mathcal{F}_t.$$

By Proposition 21, $\beta \in \mathcal{F}_{\alpha+}$; also $\Lambda \in \mathcal{F}_{\alpha+}$ by (15), hence

$$\Lambda \cap \{\beta < t\} \in \mathcal{F}_{\alpha+}.$$

Combining the two relations above we obtain

$$\Lambda \cap \{\beta < t\} \in \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} = \mathcal{F}_{(\alpha \wedge t)+} = \mathcal{E}_t.$$

Hence $\Lambda \in \mathcal{E}_{\beta+}$ by definition.

Conversely, let $\Lambda \in \mathcal{E}_{\beta+}$, then by definition

(37) $$\forall t: \ \forall u: \ \Lambda \cap \{\beta < t\} \cap \{(\alpha \wedge t) < u\} \in \mathcal{F}_u.$$

This reduces to the following two relations:

(37′) $$\forall t < u: \ \Lambda \cap \{\beta < t\} \in \mathcal{F}_u;$$

(37″) $$\forall t \geqq u: \ \Lambda \cap \{\beta < t\} \cap \{\alpha < u\} \in \mathcal{F}_u.$$

By (37′) and the optionality of α, we have

$$\forall t < u: \quad \Lambda \cap \{\beta < t\} \cap \{\alpha < u\} \in \mathcal{F}_u.$$

Combining this with (37″), and letting $t \uparrow \infty$, we obtain

$$\forall u: \quad \Lambda \cap \{\beta < \infty\} \cap \{\alpha < u\} \in \mathcal{F}_u. \tag{38}$$

Since $\Lambda \in \mathcal{E}_{\beta+}$, we have also by (32+):

$$\Lambda \cap \{\beta = +\infty\} \in \mathcal{E}_{+\infty} = \mathcal{F}_{\alpha+}$$

and consequently

$$\forall u: \quad \Lambda \cap \{\beta = +\infty\} \cap \{\alpha < u\} \in \mathcal{F}_u.$$

Combining this with (38), we obtain

$$\forall u: \quad \Lambda \cap \{\alpha < u\} \in \mathcal{F}_u. \tag{39}$$

Finally, since

$$\Lambda \cap \{(\alpha \wedge \beta) < u\} = [\Lambda \cap \{\alpha < u\}] \cup [\Lambda \cap \{\beta < u\}], \tag{40}$$

and both members of the union above belong to $\mathcal{F}_u$ by (39) and (37′), we conclude that the left member of (40) belongs to $\mathcal{F}_u$. Thus $\Lambda \in \mathcal{F}_{(\alpha \wedge \beta)+}$ and Proposition 22 is proved. (In the final step we may also apply (28).)

The interest of Propositions 21 and 22 lies in this: given the optional β, any optional α dominating β can be made to play the role of $+\infty$ if the new family $\{\mathcal{E}_t\}$ is used instead of $\{\mathcal{F}_t\}$. In particular, considerations of (β, α) may be reduced to that of $(\beta, +\infty)$.

Proposition 23. *If α is optional relative to $\{\mathcal{F}_t\}$ and $\beta \geqq \alpha$, then β is optional relative to $\{\mathcal{F}_t\}$ if and only if $\beta - \alpha$ is optional relative to $\{\mathcal{G}_t, t \in T^0\}$.*

Proof. If β is optional relative to $\{\mathcal{F}_t\}$, then we have by Proposition 16:

$$\forall t \in T^0: \quad \{\beta - \alpha < t\} = \{\beta < \alpha + t\} \in \mathcal{F}_{(\alpha+t)-} = \mathcal{G}_{t-}$$

Hence $\beta - \alpha$ is optional relative to $\{\mathcal{G}_t\}$. Conversely if this is true, then

$$\forall s \in T^0: \quad \{\beta - \alpha < s\} \in \mathcal{G}_s = \mathcal{F}_{\alpha+s}.$$

Since $\{\alpha < r\} = \{\alpha + s < r + s\}$, and $\alpha + s$ is strictly optional, we have by Definition 4:

$$\forall r \in T, s \in T^0: \quad \{\beta - \alpha < s\} \cap \{\alpha < r\} \in \mathcal{F}_{r+s}.$$

It follows that

$$\forall t \in T: \quad \{\beta < t\} = \{\alpha + (\beta - \alpha) < t\} = \bigcup_{\substack{r+s<t \\ r \in R,\, s \in R^0}} \{\alpha < r; \beta - \alpha < s\} \in \mathcal{F}_t.$$

Hence β is optional relative to $\{\mathcal{F}_t\}$.

An application we may deduce, e.g. Lemma 4.4 in [4; p. 167].

PROPOSITION 23.1. *For any optional α and constant $t_0 \in T$, $(t_0 - \alpha) \vee 0$ is optional relative to $\{\mathcal{G}_t, t \in T^0\}$ and $(\alpha - t_0) \vee 0$ is optional relative to $\{\mathcal{F}_{t_0+t}, t \in T^0\}$.*

PROPOSITION 24. *If α and β are both optional relative to $\{\mathcal{F}_t\}$ and $\alpha \leqq \beta$, then*

$$\mathcal{G}_{(\beta-\alpha)+} = \mathcal{F}_{\beta+}.$$

Proof. Let $\Lambda \in \mathcal{F}_{\beta+}$, then we have by Proposition 16, for every $t \in T^0$:

$$\Lambda \cap \{\beta - \alpha < t\} = \Lambda \cap \{\beta < \alpha + t\} \in \mathcal{F}_{(\alpha+t)-} \subset \mathcal{F}_{\alpha+t} = \mathcal{G}_t.$$

Thus $\Lambda \in \mathcal{G}_{(\beta-\alpha)+}$ since $\beta - \alpha$ is optional relative to $\{\mathcal{G}_t\}$ by Proposition 23. Conversely, let $\Lambda \in \mathcal{G}_{(\beta-\alpha)+}$, then we have for every $t \in T$:

$$\Lambda \cap \{\beta < t\} = \bigcup_{\substack{r+s<t \\ r \in R,\ s \in R^0}} [\Lambda \cap \{\beta - \alpha < s\} \cap \{\alpha < r\}]$$

As in the preceding proof, the last-written union belongs to $\mathcal{F}_t$ and so $\Lambda \in \mathcal{F}_{\beta+}$.

3. Progressive and natural Borel measurability. Let X be a space, $\mathcal{A}$ a Borel field on X. For each $t \in T$, let $x_t \colon \omega \to x_t(\omega)$ be a function on Ω to X. We write also $x(t, \omega)$ for $x_t(\omega)$. For each t, it is clear that $x_t^{-1}(\mathcal{A})$ is a B.F. on Ω. Let

$$\mathcal{F}_t^0 = \bigvee_{s \leqq t} x_s^{-1}(\mathcal{A}).$$

Then $\mathcal{F}_t^0$ is the B.F. generated by all x_s with $s \leqq t$, and $\{\mathcal{F}_t^0, t \in T\}$ is a nondecreasing family of B.F.'s on Ω.

Given $\{x_t\}$, the family of B.F.'s $\{\mathcal{F}_t\}$ on Ω is said to be *adapted to* $\{x_t\}$ iff it is nondecreasing and $x_t \in \mathcal{F}_t$ for each t. The family $\{\mathcal{F}_t^0\}$ defined above is adapted and is minimal in the sense that for any adapted family $\{\mathcal{F}_t\}$ we have $\mathcal{F}_t^0 \subset \mathcal{F}_t$ for each t.

Definition 5. The family $\{\mathcal{F}_t^0\}$ is called the *natural family* of Borel fields for $\{x_t\}$.

If α is a function on Ω to T, then the function $\omega \to x(\alpha(\omega), \omega)$ will sometimes be denoted by x_α. In particular wo shall write:

$$x_{\alpha \wedge t} \colon \quad \omega \to x(\alpha(\omega) \wedge t, \omega),$$
$$x_{\alpha+t} \colon \quad \omega \to x(\alpha(\omega) + t, \omega).$$

Note that α is supposed finite here since $x_{\pm\infty}$ have not been defined.

Let θ be an element alien to X: $\theta \notin X$. Put $X_0 = X \cup \{\theta\}$ and define $\mathcal{A}_\theta$ to be the Borel field generated by $\mathcal{A}$ and the singleton $\{\theta\}$. Now for given $\{x_t\}$ and α define:

$$\forall t \in T: \quad x_t^- = \begin{cases} x_t & \text{on } \{t < \alpha\}, \\ \theta & \text{on } \{t \geqq \alpha\}. \end{cases} \tag{42}$$

Let $\mathfrak{F}^-$ be the B. F. generated by $\{x_t^-, t \in T\}$. Recall the definition of $\mathfrak{F}_{\alpha-}{}^0$ from Definitions 1 and 5.

PROPOSITION 25. *We have* $\mathfrak{F}^- = \mathfrak{F}_{\alpha-}{}^0$.

Proof. For each t, we have

$$\{x_t^- = \theta\} = \{\alpha \leqq t\}, \tag{43}$$

hence $\alpha \in \mathfrak{F}^-$. For any A in $\mathcal{A}$ and $s \leqq t$, we have

$$\{x_s \in A; t < \alpha\} = \{x_s^- \in A\}, \tag{44}$$

from which we deduce the more general relation

$$\mathfrak{F}_t{}^0 \cap \{t < \alpha\} \subset \mathfrak{F}^-.$$

Hence each generating set of $\mathfrak{F}_{\alpha-}{}^0$ belongs to $\mathfrak{F}^-$ and so $\mathfrak{F}_{\alpha-}{}^0 \subset \mathfrak{F}^-$. Conversely, we see from (43) that

$$\{x_t^- = \theta\} \in \mathfrak{F}\{\alpha\} \subset \mathfrak{F}_{\alpha-}{}^0,$$

by (5), and from (44) that

$$\{x_t^- \in A\} \in \mathfrak{F}_{\alpha-}{}^0.$$

It follows from the definition of $\mathcal{A}_\theta$ that $x_t^- \in \mathfrak{F}_{\alpha-}{}^0$ so that $\mathfrak{F}^- \subset \mathfrak{F}_{\alpha-}{}^0$. Proposition 25 is proved.

Remark. If we replace $\{t < \alpha\}$ and $\{t \geqq \alpha\}$ in (42) respectively by $\{t \leqq \alpha\}$ and $\{t > \alpha\}$, call the resulting function x_t^*, and $\mathfrak{F}^*$ the B. F. generated by $\{x_t^*, t \in T\}$, then we have

$$\mathfrak{F}_{\alpha-}{}^0 = \mathfrak{F}^- \subset \mathfrak{F}^* = \mathfrak{F}_{\alpha*}{}^0$$

where the last B. F. is defined in the Remark after Definition 2, and the "$\subset$" above cannot be replaced by "$=$" in general (Example 6 in § 5). A similar B. F., that generated by $\{x_{\alpha \wedge t}, t \in T\}$, has been used in very special cases such as Brownian motion to play the role of $\mathfrak{F}_{\alpha-}{}^0$. However, this new field may not contain α (Example 4 in § 5) but must contain x_α, which need not be contained in $\mathfrak{F}_{\alpha-}$, nor even in $\mathfrak{F}_{\alpha+}$ for optional α (Example

5 in § 5). Indeed, we shall proceed to find a condition under which $x_\alpha \in \mathfrak{F}_{\alpha+}$ for every optional α.

As a matter of general terminology, if X_i is a set and $\mathfrak{F}_{X_i}$ is a B. F. on X_i for $i=1,2$, we shall write $f \in \mathfrak{F}_{X_1}/\mathfrak{F}_{X_2}$ iff $f^{-1}(\mathfrak{F}_{X_2}) \subset \mathfrak{F}_{X_1}$. (Thus our earlier notation e.g., "$\alpha \in \mathfrak{F}_{\alpha-}$" is an abbreviation for "$\alpha \in \mathfrak{F}_{\alpha-}/\mathcal{B}$.") The product space $X_1 \times X_2$ and product field $\mathfrak{F}_{X_1} \times \mathfrak{F}_{X_2}$ are defined as usual (see [5; pp. 150 ff.]). In this connection let us record a well-known result.

PROPOSITION 26. *Let $(X_i, \mathfrak{F}_{X_i})$, $i=1,2,3$, be three pairs of space-fields. Suppose that*

$$f \in \mathfrak{F}_{X_1} \times \mathfrak{F}_{X_2}/\mathfrak{F}_{X_3}, \quad \phi_1 \in \mathfrak{F}_{X_1} \times \mathfrak{F}_{X_2}/\mathfrak{F}_{X_1}, \quad \phi_2 \in \mathfrak{F}_{X_1} \times \mathfrak{F}_{X_2}/\mathfrak{F}_{X_2},$$

and g is the function

$$g:\ (\xi_1, \xi_2) \to f(\phi_1(\xi_1, \xi_2), \phi_2(\xi_1, \xi_2)).$$

Then $g \in \mathfrak{F}_{X_1} \times \mathfrak{F}_{X_2}/\mathfrak{F}_{X_3}$.

Now consider the pairs

$$(T, \mathcal{B}), \qquad (\Omega, \mathfrak{F}), \qquad (X, \mathfrak{A})$$

where $\mathfrak{F}$ is an arbitrary B. F. on Ω and the other symbols have been introduced before. Suppose that for each $t \in T$, we have $x_t \in \mathfrak{F}/\mathfrak{A}$, then $\mathfrak{F}_t{}^0 \subset \mathfrak{F}$ and so $\mathfrak{F}_{+\infty}{}^0 \subset \mathfrak{F}$. In the usual language $\{x_t, t \in T\}$ is a family of measurable functions on the measurable space $(\Omega, \mathfrak{F})$. The family $\{x_t\}$ is called *Borel measurable* iff the function $x(\cdot,\cdot) \in \mathcal{B} \times \mathfrak{F}/\mathfrak{A}$. The following definition is more stringent, and the rest of the section is devoted to developing its main consequences. For the applicability of the new concept see the following section.

Definition 6. The family $\{x_t\}$ is said to be *progressively Borel measurable relative to the adapted family* $\{\mathfrak{F}_t\}$ iff we have

$$(45) \qquad \forall a \in T, A \in \mathfrak{A}:\ \{(t,\omega) : t < a, x(t,\omega) \in A\} \in \mathcal{B}_a \times \mathfrak{F}_a;$$

it is said to be *naturally Borel measurable* iff it is progressively Borel measurable relative to its natural family.

In symbols, (45) may be written as

$$(45') \qquad x\,|_{T_{a-}\times\Omega} \in \mathcal{B}_a \times \mathfrak{F}_a/\mathfrak{A}$$

where $T_{a-} = (-\infty, a)$ and $x\,|_{T_{a-}\times\Omega}$ is the restriction of $x(\cdot,\cdot)$ to $T_{a-} \times \Omega$. Let us recall that $T_a = (-\infty, a]$, $\mathcal{B}_a = T_a \cap \mathcal{B}$.

It follows from the next proposition that progressive measurability relative to $\{\mathcal{F}_t\}$ or $\{\mathcal{F}_{t-}\}$ or $\{\mathcal{F}_{t+}\}$ is the same concept.

PROPOSITION 27. *An equivalent definition of progressive measurability is obtained if we replace $\mathcal{F}_a$ in* (45) *by $\mathcal{F}_{a-}$ or $\mathcal{F}_{a+}$; or if we replace the right member of* (45) *by*

$$\bigwedge_{\delta>0} [\mathcal{B}_{a+\delta} \times \mathcal{F}_{a+\delta}] = \bigwedge_{\delta>0} [\mathcal{B}_a \times \mathcal{F}_{a+\delta}] \tag{46}$$

or if we do this and also replace "$t < a$" in (45) *by "$t \leq a$."*

The proof is similar to that of Proposition 1 or 5. The equation (46) is not quite trivial but its proof will be omitted. It is not known if in general the B. F. in (46) coincides with $\mathcal{B}_a \times \mathcal{F}_{a+}$.[4]

For an arbitrary index set the product space X^I and product field $\mathcal{A}^I$ are defined in the usual way ([5; p. 158]); and the function ϕ on X^I to X will be called Borel measurable iff $\phi \in \mathcal{A}^I/\mathcal{A}$. (For certain topological B. F.'s $\mathcal{A}$ such a function is called a Baire function.)

PROPOSITION 28. *Let $\{x_t^{(i)}, t \in T\}$, $i \in I$, be a collection of families, progressively Borel measurable relative to the same adapted $\{\mathcal{F}_t, t \in T\}$ and let ϕ be a Borel measurable function on X^I to X:*

$$\phi : (\xi^{(i)}, i \in I) \to \phi(\xi^{(i)}, i \in I).$$

Then the family $\{\Phi_t, t \in T\}$, where Φ is the function on $\mathcal{B} \times \mathcal{F}$ to $\mathcal{A}$:

$$\Phi : (t, \omega) \to \phi(x^{(i)}(t, \omega), i \in I),$$

is progressively Borel measurable relative to $\{\mathcal{F}_t\}$.

This proposition, like Proposition 26, is an easy analogue of the classical result to the effect that "a Borel measurable function of Borel measurable functions is Borel measurable."

PROPOSITION 29. *Let $\{x_t\}$ be progressively Borel measurable relative to $\{\mathcal{F}_t\}$, and ϕ be a function on Ω to T_t such that $\phi \in \mathcal{F}_{t+}/\mathcal{B}_t$. If x_ϕ denotes the function*

$$x_\phi : \omega \to x(\phi(\omega), \omega),$$

then $x_\phi \in \mathcal{F}_{t+}/\mathcal{A}$.

Proof. Without using one of the equivalent definitions in Proposition 27, let us first suppose that ϕ is on Ω to T_{t-}. Consider the three pairs

$$(T_{t-}, \mathcal{B}_t), \qquad (\Omega, \mathcal{F}_{t+}), \qquad (X, \mathcal{A}),$$

---

[4] However, P. A. Meyer has proved a result which implies the truth of this if the product fields are augmented by all null sets of a produce measure on $\mathcal{B} \times \mathcal{F}$.

and the three functions: $f \equiv x\,|_{T_t \times \Omega}$, $\phi_1(t,\omega) \equiv \phi(\omega)$, $\phi_2(t,\omega) \equiv \omega$. By (45′), $f \in \mathcal{B}_t \times \mathcal{F}_{t+}/\mathcal{A}$. Hence an application of Proposition 26 yields $x_\phi \in \mathcal{B}_t \times \mathcal{F}_{t+}/\mathcal{A}$. Since x_ϕ is a function of ω alone this reduces to $x_\phi \in \mathcal{F}_{t+}/\mathcal{A}$. Now if ϕ is on Ω to T_t, we replace t by $t+\delta$ in the above for $\delta > 0$ and we have by what has just been proved: $x_\phi \in \mathcal{F}_{t+\delta}/\mathcal{A}$. This being true for every $\delta > 0$, we obtain $x_\phi \in \mathcal{F}_{t+}/\mathcal{A}$ as asserted.

PROPOSITION 30. *Let $\{x_t\}$ be progressively Borel measurable relative to $\{\mathcal{F}_t\}$, and α be finite-valued and optional relative to $\{\mathcal{F}_t\}$. Then we have*

$$\forall t \in T: \quad x_{\alpha\wedge t} \in \mathcal{F}_{(\alpha\wedge t)+} \tag{47}$$

where the "+" may be omitted if α is strictly optional; and

$$x_\alpha \in \mathcal{F}_{\alpha+}. \tag{48}$$

Proof. Since $\alpha \wedge t$ is optional, we have $\alpha \wedge t \in \mathcal{F}_{(\alpha\wedge t)+}$; since also $\alpha \wedge t \leq t$, we may apply Proposition 29 to obtain

$$\forall t \in T: \quad x_{\alpha\wedge t} \in \mathcal{F}_{t+}.$$

It follows that for any $A \in \mathcal{A}$ and $s \in T$, we have

$$\begin{aligned}&\{x_{\alpha\wedge t} \in A\} \cap \{\alpha < s\} \\ &\qquad = \{x_{\alpha\wedge(s\wedge t)} \in A\} \cap \{\alpha < s\} \in \mathcal{F}_{(s\wedge t)+} \vee \mathcal{F}_{s+} = \mathcal{F}_{s+}.\end{aligned}$$

Thus $x_{\alpha\wedge t} \in \mathcal{F}_{\alpha+}$ by Definition 4, and consequently by (28):

$$x_{\alpha\wedge t} \in \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} = \mathcal{F}_{(\alpha\wedge t)+}.$$

Hence (47) is proved and the case of strict optionality is similar. Furthermore, we have by (47) and (32+) (note that since α is finite here, (32+) follows simply from (33)):

$$x_\alpha = \lim_{t\uparrow\infty} x_{\alpha\wedge t} \in \bigvee_t \mathcal{F}_{(\alpha\wedge t)+} = \mathcal{F}_{\alpha+}.$$

Proposition 30 is proved.

Let $\alpha(\cdot,\cdot)$ be a function on $T \times \Omega$ to T with the following properties:

(i) $\forall t \in T$: $\alpha_t(\cdot) = \alpha(t,\cdot)$ is optional relative to $\{\mathcal{F}_s, s \in T\}$;

(ii) $\forall \omega \in \Omega$: $\alpha(\cdot,\omega)$ is nondecreasing and right continuous on T.

The family $\{\mathcal{F}_{\alpha_t+}, t \in T\}$ is then nondecreasing by (15), and adapted to $\{x_{\alpha_t}, t \in T\}$ by (48), where x_{α_t} is the function below:

$$x_{\alpha_t}: \omega \to x(\alpha(t,\omega),\omega).$$

Examples are, for an optional $\alpha(\cdot)$:

$$\alpha(t,\omega) \equiv \alpha(\omega) \wedge t, \qquad \alpha(\omega) \vee t, \qquad \alpha(\omega) + t,$$

the last for $t \in [0, \infty)$.

PROPOSITION 31. *If $\{x_t\}$ is progressively Borel measurable relative to $\{\mathfrak{F}_t\}$, then so is $\{x_{\alpha_t}\}$ relative to $\{\mathfrak{F}_{\alpha_{t+}}\}$.*

Proof. We prove first the following lemma which is the particular case of the proposition for $x(t,\omega) \equiv t$.

LEMMA. *The family $\{\alpha_t, t \in T\}$ (of functions on Ω to T, indexed by T) is progressively Borel measurable relative to $\{\mathfrak{F}_{\alpha_{t+}}\}$.*

Proof of the Lemma. Let $a \in T$, $c \in T$, then it follows from the hypotheses in (ii) that

$$\{(t,\omega) : t < a, \alpha(t,\omega) < c\} = \bigcup_{r \in R \cap (-\infty, a)} \{(t,\omega) : t < r, \alpha(r,\omega) < c\}. \tag{49}$$

Since $r < a$, we have $\alpha_r \leqq \alpha_a$ by (ii), and so the set in $\{\ \}$ on the right side of (49) may be written as $T_{r-} \times F$ where

$$F = \{\omega : [\alpha_r(\omega) \wedge \alpha_a(\omega)] < c\}.$$

Since $F \in \mathfrak{F}_{\alpha_{a+}}$ by (5) and (15), we see that the set on the left side of (49) belongs to $\mathcal{B}_a \times \mathfrak{F}_{\alpha_{a+}}$. Hence $\{\alpha_t\}$ is progressively Borel measurable relative to $\{\mathfrak{F}_{\alpha_{t+}}\}$ by definition. The Lemma is proved.

Next, we prove that for each $\Gamma \in \mathcal{B} \times \mathfrak{F}_{\alpha_{a+}}$, we have

$$\{(t,\omega) : t < a, (\alpha(t,\omega),\omega) \in \Gamma\} \in \mathcal{B}_a \times \mathfrak{F}_{\alpha_{a+}}. \tag{50}$$

It is sufficient to prove (50) for Γ of the form $B \times F$ where $B \in \mathcal{B}$ $F \in \mathfrak{F}_{\alpha_{a+}}$. For such a set the left member of (50) reduces to

$$(T_{a-} \times F) \cap \{(t,\omega) : t < a; \alpha(t,\omega) \in B\}. \tag{51}$$

By the lemma above, the set in $\{\ \}$ in (51) belongs to $\mathcal{B}_a \times \mathfrak{F}_{\alpha_{a+}}$; since $T_{a-} \times F$ also belongs to this field, (50) follows.

Now let $A \in \mathcal{A}$, and define two subsets Γ_1 and Γ_2 of $T \times \Omega$ as follows:

$$\Gamma_1 = \{(s,\omega) : s < \alpha(a,\omega); x(s,\omega) \in A\}, \tag{52}$$

$$\Gamma_2 = \{(s,\omega) : s = \alpha(a,\omega); x(s,\omega) \in A\}. \tag{53}$$

We have

$$\Gamma_1 = \bigcup_{r \in R} \{(s,\omega) : s < r < \alpha(a,\omega); x(s,\omega) \in A\}. \tag{54}$$

PROPOSITION 32. *If the metric space valued process $\{x_t\}$ is separably Borel measurable, $\hat{x}_.$ is separably Borel measurable. Conversely if the function $\hat{x}_.$ is separably Borel measurable there is a standard modification of the process which is natuarlly separably Borel measurable.*

Proof. (The method of proof was suggested by P. A. Meyer.) Consider the class of processes $\{x_t\}$ corresponding to separably Borel measurable functions $x(\cdot,\cdot)$ for which the corresponding function $\hat{x}_.$ is separably Borel measurable. The class Γ of functions $x(\cdot,\cdot)$ so defined contains the simple product-space functions, because for such a function $\hat{x}_.$ becomes simple. The class Γ is closed under sequential convergence and therefore contains all separably Borel measurable functions. Thus the direct half of the proposition is true. Conversely suppose that the function $\hat{x}_.$, which we shall denote by $\hat{x}(\cdot)$, is separably Borel measurable. Then the range of this function is separable. Let $S_1^n, S_2^n, \cdots$ be a disjoint partition of the closure of this range into Borel sets of diameter $< 1/n^2$, with the $(n+1)$-th partition a refinement of the n-th. Define A_{ji}^n by

$$A_{ji}^n = \{t : \hat{x}(t) \in S_j^n, i2^{-n} < t \leqq (i+1)2^{-n}\}.$$

Then for each n, $\{A_{ji}^n\}$ is a partition of T. If A_{ji}^n is not empty choose a point t_{ji}^n in it, making the choice in such a way that each t_{ji}^n is also some t_{lk}^{n+1}. Define the function ϕ_n from T to T by

$$\phi_n(t) = t_{ji}^n \text{ on } A_{ji}^n.$$

Then ϕ_n is Borel measurable and

$$x[\phi_n(t)] = x(t) \qquad \text{if } t = t_{ji}^m \text{ and } n \geqq m.$$
$$|\phi_n(t) - t| < 2^{-n}.$$

Moreover for each value of t,

$$\lim_{n\to\infty} x[\phi_n(t), \omega] = x(t, \omega) \tag{58}$$

for almost all ω. Define $x_0(t,\omega)$ as the limit on the left when the limit exists and as c otherwise, where c is some specified element of X. Then $x_0(t,\omega) = x(t,\omega)$ for $t = t_{ji}^n$ and the $x_0(t)$ process is a standard modification of the given one, determined completely by $x(t,\omega)$ for $t \in \{t_{ji}^n\}$ and by the choice of c. If $\delta > 0$ and if n is sufficiently large, $x[\phi_n(t),\omega]$ defines a process whose restriction to the interval $(-\infty, a)$ is separably Borel measurable relative to the field $\mathcal{B}_{a+\delta} \times \mathcal{F}_{a+\delta}$ where $\{\mathcal{F}_t\}$ is the natural field family for the $x_0(t)$ process. Hence the same restriction of the $x_0(t,\omega)$ process

has the same measurability property for all $\delta > 0$. The $x_0(t, \omega)$ process is therefore progressively separably Borel measurable relative to its natural field family, as was to be shown.

We have actually proved more than the proposition asserts. The new natural fields are contained in the old ones for $t \in t_{jl}{}^n$, a set dense in T. Moreover the assertions about the new process are also valid for its restriction to any interval of the form (b, ∞).

If we make the additional assumption that the range space X of the random variables is compact as well as metric, the conclusion of Proposition 32 can be strengthened. Let f be a function on T into X. For an arbitrary subset A of T, let $f[A]$ be the range of the restriction of f to A, and let $f[A]^*$ be the closure of $f[A]$ in the topology of X. For each t in T, we set

$$\begin{aligned} L_+(f, A, t) &= \bigcap_n f[[t, t + n^{-1}] \cap A]^*, \\ L_-(f, A, t) &= \bigcap_n f[[t - n^{-1}, t] \cap A]^*, \\ L(f, A, t) &= \bigcap_n f[[t - n^{-1}, t + n^{-1}] \cap A]^* \\ &\quad - L_-(f, A, t) \cup L_+(f, A, t). \end{aligned}$$

Thus $L[L_+, L_-]$ is the set of [right, left] limiting values of f on A at t. The function f is said to be *[right, left] separable at t with A as a separating set* iff

$$f(t) \in L(f, A, t) \, [L_+(f, A, t), L_-(f, A, t)].$$

It is said to be [right, left] separable iff this is so at each t in T. The process $\{x_t, t \in T\}$ taking values in X is said to be *[right, left] separable* iff there is a countable set A such that for each ω in Ω, the sample function $x(\cdot, \omega)$ is so separable with A as a separating set.[5]

PROPOSITION 33. *If X is compact metric, the standard modification described in Proposition 32 can be made separable in addition to the other stated properties.*

Proof. To prove this assertion we change the definition of $x_0(t, \omega)$ in the proof of Proposition 32. Let $\{\xi_n\}$ be a sequence dense in X and define: $v_{11}(t, \omega)$ is the smallest value of $n \geqq 1$ for which

$$\rho(\xi_1, x[\phi_n(t), \omega]) \leqq \liminf_{m \to \infty} \rho(\xi_1, x[\phi_m(t), \omega]) + 1;$$

[5] These definitions were given by Chung in unpublished lecture notes in 1962, in which he proved that every real-valued process has a standard modification which is right [left] separable.

$v_{1j}(t,\omega)$ is the smallest value of $n > v_{1,j-1}(t,\omega)$ for which

$$\rho(\xi_1, x[\phi_n(t),\omega]) \leqq \liminf_{m\to\infty} \rho(\xi_1, x[\phi_m(t),\omega]) + 1/j.$$

Then

$$\lim_{j\to\infty} \rho(\xi_1, x[\phi_{v_{1j}}(t),\omega]) = \liminf_{m\to\infty} \rho(\xi_1, x[\phi_m(t),\omega])$$

and it is clear that if $\delta > 0$ the restriction of v_{1j} to $T_{a-}\times\Omega$ belongs to $\mathcal{B}_{a+\delta}\times\mathcal{F}_{a+\delta}$ for sufficiently large j. In general choose $\{v_{kj}(t,\omega), j \geqq 1\}$ a subsequence of $\{v_{k-1,j}(t,\omega), j\geqq 1\}$ in such a way that the restriction of v_{kj} to $T_{a-}\times\Omega$ belongs to $\mathcal{B}_{a+\delta}\times\mathcal{F}_{a+\delta}$ for sufficiently large j, whenever $\delta > 0$, and that

$$\lim_{j\to\infty} \rho(\xi_k, x[\phi_{v_{kj}}(t),\omega]) = \liminf_{j\to\infty} \rho(\xi_k, x[\phi_{v_{k-1,j}}(t),\omega])$$

Then

$$\lim_{j\to\infty} \rho(\xi_k, x[\phi_{v_{jj}}(t),\omega])$$

exists for all k, t, ω. Hence $\lim_{j\to\infty} x[\phi_{v_{jj}}(t),\omega]$ exist for all (t,ω) and we define $x_0(t)$ as this limit. Clearly the x_{0t} process is a standard modification of the x_t process with all the properties required in the proposition. The separating sequence is the sequence $\{t_{ji}{}^n\}$ defined in the proof of Proposition 32. Finally we note that the x_{0t} process is 'right separable' as defined above, if $\phi_n(t) \geqq t$. This inequality is not necessarily satisfied as we have defined ϕ_n, but the definition can be modified to achieve this inequality as follows. Define $A_{ji}{}^n$ as in the proof of Proposition 32. If $A_{ji}{}^n$ contains its supremum let $t_{ji}{}^n$ be this supremum and define $\phi_n(t) = t_{ji}{}^n$ on $A_{ji}{}^n$ as before. Otherwise let $\{t_{jik}{}^n, k \geqq 1\}$ be a monotone sequence in $A_{ji}{}^n$ with limit equal to this supremum and define

$$\phi_n(t) = t_{jik}{}^n \text{ on } A_{ji}{}^n \cap (t_{ji,k-1}{}^n, t_{jik}{}^n], \quad k \geqq 1 \text{ where } t_{ji0}{}^n = i2^{-n}.$$

As a complement we shall consider functions from a measure space to a metric space. The measure, say ν will be supposed complete. A function from the measure space to a metric space will be called ν-measurable iff it is the ν-almost everywhere limit of a sequence of simple functions. The range of the function is then ν-almost separable, that is, the restriction of the function to the complement of some set of ν-measure 0 is separable. If a function is separably Borel measurable it is ν-measurable and conversely a ν-measurable function coincides ν-almost everywhere with some separably Borel measurable function. A Borel measurable function is ν-measurable if and only if it is ν-almost separably valued. The same assertions are true if Borel

measurability is defined not using the domain of ν but any Borel subfield whose completion under ν yields the given domain of ν. In the following, μ is any Lebesgue-Stieltjes measure on $(-\infty, +\infty)$, that is, a completed measure of Borel sets, finite for compact sets. Let $\mathcal{B}^*$ be the domain of μ. The completed product measure $\mu \times P$ is defined on an extension of $\mathcal{B}^* \times \mathfrak{F}$. If $\{x_t\}$ is a stochastic process with state space X, as usual, it will be called μ-measurable iff $x(\cdot,\cdot)$ is $(\mu \times P)$-measurable.

PROPOSITION 34. *Suppose that X is metric. Then if the process $\{x_t\}$ has a μ-measurable standard modification the function $\hat{x}_.$ is μ-measurable. Conversely, if $\hat{x}_.$ is μ-measurable, there is a naturally μ-measurable standard modification.*

This proposition is due to Yukiyosi Kawada [6] aside from the 'naturally.' It is easily deduced from Proposition 32 by exploiting the relations between Borel measurable functions and functions measurable with respect to the domain of a measure, or can be deduced directly, as Kawada did.

PROPOSITION 35. *If X is compact metric, the standard modification described in Proposition 34 can be made separable (or even right separable) in addition to the other stated properties.*

Proof. The proof of Proposition 33 together with the known relations between Borel and μ-measurability yields a sequence $\{v_{jj}\}$ such that $\lim_{j\to\infty} x[\phi_{v_{jj}}(t),\omega]$ exists for all (t,ω) and, if t is not in an exceptional Borel set B of μ-measure 0, the limit is $x(t,\omega)$ with probability 1. If t is not in B define $x_0(t,\omega)$ as the above limit. If t is in B any definition of $x_0(t,\omega)$ making $x_{0t} = x_t$ with probability 1 will yield an x_{0t} process which is a progressively μ-measurable standard modification of the x_t process. The standard separability argument yields a choice of x_{0t} making the process separable, or right separable if desired. (In the latter case ϕ_n must be chosen to make $\phi_n(t) \geqq t$.)

5. Examples. The following simple process may be used to furnish several examples alluded to in preceding sections. It will be described in an informal way using the terminology of [1], to which we refer for rigorous details.

There is a homogeneous Markov chain [6] $\{x_t, t \geqq 0\}$ on a probability space $(\Omega, \mathfrak{F}, P)$ with three states $\{0,1,2\}$ having the following properties. The

[6] Also called "Markov chain with stationary transition probabilities."

mean sojourn time in each state is 1. The initial state is 0, upon exit from which there is a jump to state 1 or state 2 with probability 1/2 each. Upon exit from 1 there is a jump to 2, and vice versa. Each sample function takes the value 0 in a proper interval beginning at 0 and takes the values 1 and 2 (whichever comes first) thereafter in alternate intervals extending to infinity. In one version x_+ of the process every sample function is right continuous, in another version x_-, it is left continuous. Let α be the first entrance time into the state 1 and β that into the state 2. Both have the same density function, and $\alpha \wedge \beta = \gamma$ is the exit time from the state 0, with the density $e^{-t}\,dt$.

Let $\{\mathfrak{F}_t, t \geqq 0\}$ be the natural family of $\{x_+(t), t \geqq 0\}$ or $\{x_-(t), t \geqq 0\}$; for each t let $\mathfrak{F}_t^*$ be the smallest Borel field containing $\mathfrak{F}_t$ and all sets of probability zero. Where no version of the above process is specified below, either x_+ or x_- will do.

Example 1. Relative to the family $\{\mathfrak{F}_t, t \geqq 0\}$, α is optional but not strictly so. Since $P\{\alpha = t\} = 0$ for each t, α is strictly optional relative to $\{\mathfrak{F}_t^*, t \geqq 0\}$.

Example 2. Relative to $\{\mathfrak{F}_t^*\}$, we have $\mathfrak{F}_{\alpha*} = \mathfrak{F}\{\alpha\}$, the Borel field generated by α alone. But

$$x(\alpha + 0) = \lim_{t \downarrow \alpha(\omega)} x(t, \omega) \in \mathfrak{F}_{\alpha+} = \mathfrak{F}_\alpha$$

and $P\{x(\alpha + 0) = 1\} = P\{x(\alpha + 0) = 2\} = 1/2$. Thus α is strictly optional but $\mathfrak{F}_{\alpha*} \neq \mathfrak{F}_\alpha$.

Example 3. Relative to $\{\mathfrak{F}_t\}$ or $\{\mathfrak{F}_t^*\}$, we have

$$\mathfrak{F}_{\gamma-} = \mathfrak{F}_{(\alpha\wedge\beta)-} = \mathfrak{F}\{\gamma\}$$

or the augmentation of $\mathfrak{F}\{\gamma\}$ by all sets of probability zero. The set $\{\alpha = \beta\}$ is empty; the set $\{\alpha < \beta\}$ has probability 1/2 and is independent of the random variable γ. We have by (23),

$$\{\alpha < \beta\} \in \mathfrak{F}_{\beta-}, \qquad \{\alpha < \beta\} = \Omega - \{\beta < \alpha\} \in \mathfrak{F}_{\alpha-},$$

hence

$$\{\alpha < \beta\} \in \mathfrak{F}_{\alpha-} \wedge \mathfrak{F}_{\beta-}$$

but

$$\{\alpha < \beta\} \notin \mathfrak{F}_{(\alpha\wedge\beta)-}.$$

Example 4. For the x_- version, $x_{\alpha\wedge t} \equiv 0$ for every $t \geqq 0$. Hence

$$\alpha \notin \mathfrak{F}\{x_{\alpha\wedge t}, t \geqq 0\}.$$

Example 5. Let $\Delta \notin \mathfrak{F}_{\gamma+0}$, e. g., $\Delta = \{\alpha > 1\}$; and define

$$\tilde{x}(t,\omega) = \begin{cases} x(t,\omega), & \text{if } t \neq \gamma(\omega); \\ x(\gamma(\omega)-0,\omega), & \text{if } t = \gamma(\omega), \omega \in \Delta; \\ x(\gamma(\omega)+0,\omega), & \text{if } t = \gamma(\omega), \omega \in \Omega - \Delta. \end{cases}$$

Then $\{\tilde{x}_t\}$ is a (separable) standard modification of $\{x_t\}$ and so has the same augmented natural family $\{\mathfrak{F}_t^*\}$ as $\{x_t\}$. We have $\tilde{x}_\gamma \notin \mathfrak{F}_{\gamma+0}^*$ since

$$\{\tilde{x}_\gamma = 0\} = \Delta \notin \mathfrak{F}_{\gamma+0}^*.$$

Example 6. For the x_+ process, the set $\{\alpha < \beta\}$ belongs to $\mathfrak{F}_{\alpha+}{}^0$ but not to $\mathfrak{F}_{\alpha-}{}^0$. The process $\{x_t^*\}$ is not separable, but an obvious discrete parameter analogue serves the same purpose and eliminates the question of separability.

Example 7. For the minimal chain studied in [1; § II. 19] and [2], we have, if τ_n is the n-th jump,

$$\bigvee_{n=1}^{\infty} \tau_n = \lim \tau_n = \tau$$

where τ is the "first infinity." The random variable $x(\tau+0)$ does not belong to $\bigvee_{n=1}^{\infty} \mathfrak{F}_{\tau_n+}$ but belongs to $\mathfrak{F}_{\tau+}$. See Theorem 4.4 of [2].

Acknowledgment. A number of results similar to those in this paper were obtained independently by P. A. Meyer and will appear in his forthcoming book.

STANFORD UNIVERSITY
AND
UNIVERSITY OF ILLINOIS.

REFERENCES.

[1] K. L. Chung, *Markov chains with stationary transition probabilities*, Springer, 1960.

[2] ———, "On the boundary theory for Markov chains," *Acta Mathematica*, vol. 110 (1963), pp. 19-77.

[3] J. L. Doob, *Stochastic processes*, Wiley, 1953.

[4] E. B. Dynkin, *Foundations of the theory of Markov processes* (in Russian), Moscow, 1963.

[5] P. R. Halmos, *Measure theory*, Van Nostrand, 1950.

[6] Y. Kawada, "On measurable stochastic process," *Proceedings of the Physics and Mathematics Society of Japan*, vol. 23 (1941), pp. 511-527.

HOLOMORPHIC IMBEDDINGS OF SYMMETRIC DOMAINS INTO A SIEGEL SPACE.[1]

By Ichiro Satake.

The purpose of this paper is to answer the following question raised recently by Kuga in connection with his study on (algebraic) families of abelian varieties [5]:

For a symmetric (bounded) domain $\mathfrak{D}$, determine all holomorphic isometries ρ *of* $\mathfrak{D}$ *into a Siegel space* (i.e. the space of all symmetric $n \times n$ complex matrices Z with $\operatorname{Im} Z > 0$) *such that the image* $\rho(\mathfrak{D})$ *is totally geodesic.*

If $\mathfrak{g}$ denotes the Lie algebra of the group of all analytic automorphisms of $\mathfrak{D}$, the problem is equivalent to determining all faithful representations ρ of $\mathfrak{g}$ into the Lie algebra of $Sp(2n, \boldsymbol{R})$ (symplectic group) satisfying a certain analyticity condition (H_1) (see 1.1). For convenience in applications, we shall actually consider this problem in a slightly generalized form, allowing compact components in $\mathfrak{g}$ and removing the condition of faithfulness of ρ. After making some preparations indispensable for the formulations of the problem and the results, in § 1, we shall make, in § 2, reductions of the problem to the case where $\mathfrak{g}$ is a simple (non-compact) Lie algebra corresponding to an irreducible symmetric domain and where the representation ρ is an absolutely irreducible representation of $\mathfrak{g}$ into the Lie algebra of $SU(p,q)$ (special unitary group of a hermitian form of signature (p,q)) satisfying a stronger condition (H_2). The determination of all such absolutely irreducible representations will be given in § 3. It will turn out, as has been expected, that there are only few such representations.

Notations. For a field k, $M_{p,q}(k)$ denotes the set of all $p \times q$ matrices with entries in k, and $M_p(k)$ stands for $M_{p,p}(k)$. For $X_i \in M_{n_i}(k)$ $(1 \leqq i \leqq r)$, we denote by $\operatorname{diag.}(X_1, \cdots, X_r)$ the matrix

$$\begin{pmatrix} X_1 & & 0 \\ & \ddots & \\ 0 & & X_r \end{pmatrix}$$

Received June 26, 1964.

[1] The results of this paper have been reported at the Conference on Complex Analysis held at the University of Minnesota in March, 1964 ([7]).

of degree $\sum_{i=1}^{r} n_i$. 1_n means, as usual, the identity matrix of degree n and, for a vector space V, 1_V means the identity transformation of V onto itself, both being sometimes abbreviated simply as 1. For a mapping f defined on V and for a (linear) subspace V_1 of V, the restriction of f on V_1 will be denoted as $f \mid V_1$. In particular, for a hermitian (or skew-hermitian) form F on V (or, more precisely, on $V \times V$), $F \mid V_1$ stands for the hermitian (or skew-hermitian) form on V_1 obtained by restricting it on $V_1 \times V_1$. For a hermitian form (or matrix) F, the notation $F > 0$ means that F is positive-definite.

§ 1. Exposition of the problem.

1.1. Let $\mathfrak{D}$ and $\mathfrak{D}'$ be symmetric domains (i.e. symmetric hermitian spaces of non-compact type).[2] We shall consider the problem of determining all isometries ρ of $\mathfrak{D}$ into $\mathfrak{D}'$ satisfying the following conditions:

(i) $\rho(\mathfrak{D})$ is totally geodesic in $\mathfrak{D}'$;

(ii) ρ is holomorphic.

Actually we shall solve this problem in the case where $\mathfrak{D}'$ is a Siegel space (i.e. an irreducible symmetric domain of type (III)).

Let G (resp. G') be the connected component of the identity of the group of all analytic automorphisms (or, what is the same, that of the group of all isometries) of $\mathfrak{D}$ (resp. $\mathfrak{D}'$). Then G and G' are connected semi-simple Lie groups of non-compact type, with center reduced to the identity, and $\mathfrak{D}$ and $\mathfrak{D}'$ are homogeneous spaces of G and G', respectively, the isotropy subgroups being maximal compact subgroups. It is known that any isometry ρ of $\mathfrak{D}$ into $\mathfrak{D}'$ satisfying (i) gives rise to a local isomorphism of G into G', denoted again by ρ, in the following manner. Namely, denote by τ_z the symmetry of $\mathfrak{D}$ around a point $z \in \mathfrak{D}$, and define similarly $\tau_{z'}$ for $z' \in \mathfrak{D}'$; and call G_1 the (closed) subgroup of G' formed of all products of an even number of symmetries around points in $\rho(\mathfrak{D})$. One defines a mapping ϕ from G_1 into G by setting

$$\phi(\tau_{\rho(z_1)} \cdots \tau_{\rho(z_m)}) = \tau_{z_1} \cdots \tau_{z_m} \qquad (m\colon \text{even});$$

then it can be proved that ϕ is well-defined and is a homomorphism of G_1 onto G, which is locally one-to-one. The local isomorphism $\rho\colon G \to G'$ (which is in general many-valued) is then given by ϕ^{-1}.

[2] For fundamental concepts concerning symmetric domains, see [1], [4].

Now, fix the origins $z_0 \in \mathcal{D}$ and $z_0' \in \mathcal{D}'$ once and for all and call K (resp. K') the isotropy subgroup of G (resp. G') at z_0 (resp. z_0'). We denote by $\mathfrak{g}, \mathfrak{g}', \mathfrak{k}, \mathfrak{k}', \cdots$ the Lie algebras of $G, G', K, K', \cdots$, respectively; let further

$$(0) \qquad \mathfrak{g} = \mathfrak{k} + \mathfrak{p}, \qquad \mathfrak{g}' = \mathfrak{k}' + \mathfrak{p}'$$

be the corresponding Cartan decompositions. Suppose that ρ is an isometry of $\mathcal{D}$ into $\mathcal{D}'$ satisfying (i) and such that $\rho(z_0) = z_0'$. We denote again by ρ the monomorphism of $\mathfrak{g}$ into $\mathfrak{g}'$ induced by ρ, and by τ (resp. τ') the (involutive) automorphism of $\mathfrak{g}$ (resp. $\mathfrak{g}'$) induced by τ_{z_0} (resp. $\tau_{z_0'}$); τ and τ' are nothing other than the 'Cartan involutions' of $\mathfrak{g}$ and $\mathfrak{g}'$ corresponding to the decompositions (0). Then we have $\rho \circ \tau = \tau' \circ \rho$, or, in other words,

$$(1) \qquad \rho(\mathfrak{k}) \subset \mathfrak{k}', \qquad \rho(\mathfrak{p}) \subset \mathfrak{p}'.$$

Now, since $\mathcal{D}$ and $\mathcal{D}'$ are hermitian, it is well-known that $\mathfrak{k}$ and $\mathfrak{k}'$ have non-trivial center and that there exists a uniquely determined element H_0 (resp. H_0') in the center of $\mathfrak{k}$ (resp. $\mathfrak{k}'$) such that $\mathrm{ad}(H_0)$ (resp. $\mathrm{ad}(H_0')$) induces on $\mathfrak{p}$ (resp. $\mathfrak{p}'$), identified in a natural manner with the tangent space to $\mathcal{D}$ (resp. $\mathcal{D}'$) at z_0 (resp. z_0'), the given complex structure. We denote the complexifications of $\mathfrak{g}, \cdots$ by $\mathfrak{g}_C, \cdots$ and put

$$\mathfrak{p}_\pm = \{X \in \mathfrak{p}_C \mid [H_0, X] = \pm iX\}$$

and define similarly $\mathfrak{p}_\pm'$. Then ρ satisfies the condition (ii) if and only if we have $\rho \circ \mathrm{ad}(H_0) = \mathrm{ad}(H_0') \circ \rho$ on $\mathfrak{p}$, or, in other words, if and only if

$$(2) \qquad \rho(\mathfrak{p}_+) \subset \mathfrak{p}_+', \qquad \rho(\mathfrak{p}_-) \subset \mathfrak{p}_-'.$$

The conditions (1) and (2), taken together, are clearly equivalent to the following condition:

$$(H_1) \qquad \rho \circ \mathrm{ad}(H_0) = \mathrm{ad}(H_0') \circ \rho.$$

(In the present case, (2) implies (1); but this will not be so, in a more general setting which we shall adopt in this paper.)

Thus, in sum, an isometry $\rho\colon \mathcal{D} \to \mathcal{D}'$ satisfying (i) and (ii) (and such that $\rho(z_0) = z_0'$) gives rise in a natural manner to a monomorphism $\rho\colon \mathfrak{g} \to \mathfrak{g}'$ satisfying the condition (H_1). It is obvious that, conversely, (for given z_0 and z_0') any monomorphism $\rho\colon \mathfrak{g} \to \mathfrak{g}'$ satisfying (H_1) comes from a uniquely determined isometry $\rho\colon \mathcal{D} = G/K \to \mathcal{D}' = G'/K'$ satisfying (i) and (ii). Therefore our problem is equivalent to *determining all monomorphisms of* $\mathfrak{g}$ *into* $\mathfrak{g}' = (III)_{p'}$ *satisfying the condition* (H_1).

We note here that the condition (H_1) is implied by the following condition:

$$(H_2) \qquad \rho(H_0) = H_0'.$$

1. 2. Let ρ_i $(i=1,2)$ be isometries of $\mathfrak{D}$ into $\mathfrak{D}'$ satisfying (i) and such that $\rho_i(z_0)=z_0'$. ρ_1 and ρ_2 will be regarded equivalent, if there exists an element $k'\in K'$ such that $\rho_2=k'\circ\rho_1$. The ρ_i's denoting, as above, also the corresponding monomorphisms of $\mathfrak{g}$ into $\mathfrak{g}'$, this condition amounts to saying that

$$\rho_2(X)=\mathrm{ad}(k')(\rho_1(X)) \qquad \text{for all } X\in\mathfrak{g}. \tag{3}$$

For convenience, we shall call any two homomorphisms ρ_1 and ρ_2 of $\mathfrak{g}$ into $\mathfrak{g}'$ satisfying this condition (k)-*equivalent*. The conditions (H_1) and (H_2) being clearly invariant under (k)-equivalence, our problem will be actually to determine all (k)-equivalence-classes of the monomorphisms of $\mathfrak{g}$ into $\mathfrak{g}'$ satisfying (H_1).

For instance, suppose a Cartan subalgebra $\mathfrak{h}$ (resp. $\mathfrak{h}'$) of $\mathfrak{k}$ (resp. $\mathfrak{k}'$) is given. Then any homomorphism satisfying (1) is (k)-equivalent to a homomorphism satisfying further the condition

$$\rho(\mathfrak{h})\subset\mathfrak{h}'. \tag{4}$$

Thus, on determining homomorphisms satisfying (H_1) (or (H_2)) up to (k)-equivalence, we may suppose, without any loss of generality, that they further satisfy the additional condition (4).

As stated in the Introduction, we shall consider the problem in a slightly more general form as follows. For simplicity, we call a semi-simple Lie algebra (over $\boldsymbol{R}$) *of herimitian type,* if all the non-compact simple components of it correspond to symmetric domains. Then it is clear that the conditions (H_1) and (H_2) as well as the notion of (k)-equivalence can be formulated, in an obvious way, also for any two semi-simple Lie algebras of hermitian type $\mathfrak{g}$ and $\mathfrak{g}'$, ($\mathfrak{k}$ and $\mathfrak{k}'$ being the maximal compact subalgebras of $\mathfrak{g}$ and $\mathfrak{g}'$, respectively). Thus our problem will be as follows:

For a given semi-ismple Lie algebra of hermitian type $\mathfrak{g}$, *determine all* (k)-*equivalence-classes of homomorphisms* ρ *of* $\mathfrak{g}$ *into* $\mathfrak{g}'=(III)_{p'}$ *satisfying the condition* (H_1).

(It will turn out that for two such homomorphisms ρ_1 and ρ_2, the (k)-equivalence coincides with the usual equivalence. See 2.4, Cor. to Th. 1.)

1. 3. Irreducible symmetric domains are classified by E. Cartan [1] into four series of classical domains $(I)_{p,q}$, $(II)_p$, $(III)_p$, $(IV)_p$ and two exceptional domains $(EIII)$, $(EVII)$. (Cf. also [4], [8].) In this paper, we use these symbols to denote both symmetric domains and the corresponding Lie

algebras. In order to settle the notations, we shall give here descriptions of the domains $(I)_{p,q}$ and $(III)_p$ in the form adapted to our purpose (cf. [9]).

$(I)_{p,q}$: Let V be a vector-space over $\boldsymbol{C}$ provided with a non-degenerate hermitian (sesqui-linear) form F of signature (p,q). (We usually suppose F to be indefinite, i.e. $p,q>0$; but the following considerations apply trivially also to the definite case.) One defines $\mathfrak{D}=\mathfrak{D}(V,F)=(I)_{p,q}$ as the space of all q-dimensional (complex) subspaces V_- of V such that $F\mid V_-$ is negative-definite; $\mathfrak{D}$ is thus an open submanifold of a complex Grassmannian manifold. For such a V_-, we have an orthogonal decomposition of V:

$$V=V_+ + V_-,$$

where $V_+=(V_-)^\perp$ is a p-dimensional subspace of V such that $F\mid V_+$ is positive-definite. Thus the points z in $\mathfrak{D}$ are in a one-to-one correspondence with the couples (V_+,V_-) with these properties. Notation: $\mathfrak{D}\ni z\leftrightarrow(V_+,V_-)$. (One can define another complex structure of $\mathfrak{D}$ in regarding it as the space of the V_+'s. This is nothing but the conjugate complex structure, and these two complex structures are mutually equivalent.) The connected component of the group of all analytic automorphisms of $\mathfrak{D}$ is given by $G=PU(V,F)$ (projective unitary group) operating on $\mathfrak{D}$ in a natural manner, and the isotropy subgroup K at $z_0\leftrightarrow(V_+^{(0)},V_-^{(0)})$ is given by the image in G of

$$U(V_+^{(0)},\ F\mid V_+^{(0)})\times U(V_-^{(0)},\ F\mid V_-^{(0)});$$

K can also be expressed in the form $PU(V,F)\cap PU(V,F_u)$, where F_u is a positive-definite hermitian form on V defined by the formula

$$F_u(x,y)=F(Tx,y)\qquad\text{for all } x,y\in V, \tag{5}$$

T denoting a linear transformation of V defined as follows:

$$T=\begin{cases}\ \ 1 & \text{on } V_+^{(0)},\\ -1 & \text{on } V_-^{(0)}.\end{cases} \tag{6}$$

(It is clear that, by this relation, the linear transformations T of V with $T^2=1$, and such that F_u defined by (5) is positive-definite, are in a one-to-one correspondence with the orthogonal decompositions $V=V_++V_-$.) If one denotes by $\mathfrak{g}$ the Lie algebra of G, the Cartan involution of $\mathfrak{g}$ at z_0 is given by the correspondence $X\to TXT$ $(X\in\mathfrak{g})$.

Now, take an orthonormal basis $(e_1,\cdots,e_{p+q})$ of V in such a way that $(e_1,\cdots,e_p)$ and $(e_{p+1},\cdots,e_{p+q})$ are orthonormal basis of $V_+^{(0)}$ and $V_-^{(0)}$, respectively. Then the matrices of the hermitian forms F and F_u are given by $1_{p,q}=\operatorname{diag.}(1_p,-1_q)$ and 1_{p+q}, respectively. To a point $z\leftrightarrow(V_+,V_-)$

of $\mathfrak{D}$, there corresponds uniquely a $p \times q$ complex matrix $Z = (z_{ij})$ with $1_q - {}^t\bar{Z}Z > 0$, by the relation that $\sum_{i=1}^{p} e_i z_{ij} + e_{p+j}$ $(1 \leqq j \leqq q)$ form a basis of V, and, in this way, $\mathfrak{D}$ is realized as a bounded domain in $\boldsymbol{C}^{pq}$. For the corresponding Lie algebras, we have the following matricial expressions:

$$\mathfrak{g} = \{ \begin{array}{cc} \overbrace{}^{p} & \overbrace{\phantom{X_{12}}}^{q} \end{array} \left(\begin{array}{cc} X_1 & X_{12} \\ {}^t\bar{X}_{12} & X_2 \end{array} \right) \begin{array}{l} \}p \\ \}q \end{array} \mid {}^t\bar{X}_i = -X_i \ (i = 1, 2), \operatorname{tr}(X_1) + \operatorname{tr}(X_2) = 0 \},$$

$$\mathfrak{k} = \{ \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \in \mathfrak{g} \},$$

$$\mathfrak{p}_+ = \{ \begin{pmatrix} 0 & X_{12} \\ 0 & 0 \end{pmatrix} \mid X_{12} \in M_{p,q}(\boldsymbol{C}) \},$$

$$\mathfrak{h} = \{ \operatorname{diag.}(\xi_1, \cdots, \xi_{p+q}) \mid \xi_i \in \sqrt{-1}\boldsymbol{R} \},$$

$$H_0 = \operatorname{diag.}\left(\frac{q}{p+q}\sqrt{-1}\,1_p, \ -\frac{p}{p+q}\sqrt{-1}\,1_q \right). \tag{7}$$

The root system of $\mathfrak{g}_C$ relative to $\mathfrak{h}_C$ is given by $\{ \xi_i - \xi_j \ (i \neq j) \}$. We shall later use thet following lemma which can easily be verified from the definitions.

LEMMA 1. *Let $\mathfrak{g} = \mathfrak{g}(V, F) = (I)_{p,q}$ and let $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ be a Cartan decomposition of it taken at $z_0 \leftrightarrow (V_+^{(0)}, V_-^{(0)})$. Let H_0 be the element in the center of $\mathfrak{k}$ giving the complex structure of $\mathfrak{p}$ and T the linear transformation of V defined by* (6). *Then we have*

$$T = -2iH_0 + \frac{p-q}{p+q} 1_V. \tag{7a}$$

1.4. $(III)_p$: Let $V_{\boldsymbol{R}}$ be a real vector-space of dimension $2p$ provided with a non-degenerate alternating bilinear form A. Call $V = V_{\boldsymbol{C}}$ the complexification of $V_{\boldsymbol{R}}$ and let $x \to \bar{x}$ be the conjugation of V relative to $V_{\boldsymbol{R}}$. Then, extending A in a natural manner to an alternating form on V, denoted again by A, we have

$$A(\bar{x}, \bar{y}) = \overline{A(x, y)} \qquad \text{for all } x, y \in V. \tag{8}$$

Now put

$$F(x, y) = iA(\bar{x}, y) \qquad \text{for all } x, y \in V. \tag{9}$$

Then one sees at once that F is a hermitian form on V of signature (p, p). One defines $\mathfrak{D} = \mathfrak{D}(V_{\boldsymbol{R}}, A) = (III)_p$ as the space of all complex structures I

on V_R such that the bilinear form $A(x, Iy)$ $(x, y \in V_R)$ is symmetric and positive-definite. Suppose that such a complex structure I is given. Extending I in a natural manner to a linear transformation of V, denoted again by I, put

$$W = \{x \in V \mid Ix = ix\}. \tag{10}$$

Then one obtains a direct decomposition

$$V = W + \bar{W}, \tag{11}$$

and sees easily that W satisfies the following properties:

$$A \mid W = 0, \qquad F \mid W > 0, \tag{12}$$

which imply, in particular, that $\bar{W}$ is orthogonal complement of W with respect to F. Conversely, given a p-dimensional (complex) subspace W of V with the properties (12), then one gets a direct decomposition (11), which determines by (10) a complex structure I on V_R satisfying the above condition. Thus $\mathfrak{D} = \mathfrak{D}(V_R, A)$ may also be regarded as the space of all p-dimensional subspaces $\bar{W}$ of V with W satisfying (12), which is a closed submanifold of $\mathfrak{D}(V, F)$. The group of all analytic automorphisms of $\mathfrak{D}$ is given by $G = PSp(V_R, A)$ (projective symplectic group), which is connected, the isotropy subgroup at $z_0 \leftrightarrow (W^{(0)}, \bar{W}^{(0)})$ being given by $K = PSp(V_R, A) \cap PU(V, F_u)$, which is isomorphic to $U(W^{(0)}, F \mid W^{(0)})/\{\pm 1\}$.

Now take an orthonormal basis $(e_1, \cdots, e_p)$ of $W^{(0)}$ with respect to F and put $e_{p+i} = \bar{e}_i$ for $1 \leqq i \leqq p$. Then $(e_1, \cdots, e_{2p})$ is an orthogonal basis of V with respect to F as described in 1.3. The matrices of the alternating form A and the conjugation: $x \to \bar{x}$ with respect to this basis are given by $\begin{pmatrix} 0 & i1_p \\ -i1_p & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1_p \\ 1_p & 0 \end{pmatrix}$, respectively. In the manner as explained in 1.3, $\mathfrak{D}$ is then realized as a bounded domain in $\boldsymbol{C}^{p(p+1)/2}$, formed of all symmetric $p \times p$ complex matrices Z with $1_p - \bar{Z}Z > 0$. We obtain also the following expressions for the Lie algebras:

$$\mathfrak{g} = \{\begin{pmatrix} X_1 & X_{12} \\ \bar{X}_{12} & \bar{X}_1 \end{pmatrix} \mid X_1, X_{12} \in M_p(\boldsymbol{C}), {}^t\bar{X}_1 = -X_1, {}^tX_{12} = X_{12}\},$$

$$\mathfrak{k} = \{\begin{pmatrix} X_1 & 0 \\ 0 & \bar{X}_1 \end{pmatrix} \in \mathfrak{g}\},$$

$$\mathfrak{p}_+ = \{\begin{pmatrix} 0 & X_{12} \\ 0 & 0 \end{pmatrix} \mid X_{12} \in M_p(\boldsymbol{C}), {}^tX_{12} = X_{12}\},$$

$$\mathfrak{h} = \{\mathrm{diag.}(\xi_1, \cdots, \xi_p, -\xi_1, \cdots, -\xi_p) \mid \xi_i \in \sqrt{-1}\,\boldsymbol{R}\},$$

$$H_0 = \mathrm{diag.}(\frac{\sqrt{-1}}{2} 1_p, \ -\frac{\sqrt{-1}}{2} 1_p).$$

The root system of $\mathfrak{g}_C$ relative to $\mathfrak{h}_C$ is given by

$$\{\pm\xi_i\pm\xi_j(i\neq j),\ \pm2\xi_i(1\leqq i\leqq p)\}.$$

The above injection $\mathfrak{D}(V_R, A)\to\mathfrak{D}(V,F)$ corresponds to the identical homomorphism of Lie algebras: $(III)_p\to(I)_{p,p}$, satisfying clearly (H_2).

1. 5. We shall define here a natural injection $\iota_{p,q}$ of $(I)_{p,q}$ into $(III)_{p+q}$ satisfying the condition (H_1).

Let (V,F) be a hermitian vector-space of signature (p,q). Call V^* the dual space of V and define a $\boldsymbol{C}$-antilinear isomorphism $V\ni x\to x^*\in V^*$ by the formula

$$F(x,y)=i\langle x^*,y\rangle \qquad \text{for all } x,y\in V. \tag{13}$$

Furthermore put $\tilde{V}=V\oplus V^*$ and extend the correspondence $x\leftrightarrow x^*$ to a conjugation σ of $\tilde{V}$ defined by

$$\sigma:\ x+y^*\to y+x^* \qquad \text{for all } x,y\in V. \tag{14}$$

One defines also a (non-degenerate) alternating for $\tilde{A}$ on $\tilde{V}$ by

$$\begin{aligned}\tilde{A}(x_1+y_1^*,x_2+y_2^*)&=\langle y_1^*,x_2\rangle-\langle y_2^*,x_1\rangle\\ &=-iF(y_1,x_2)+iF(y_2,x_1),\end{aligned} \tag{15}$$

which clearly satisfies (8) with respect to the conjugation σ. Hence we may regard $\tilde{V}$ as a complexification of a real symplectic vector-space $(\tilde{V}_R,\tilde{A})$. The corresponding hermitian form on $\tilde{V}$ (defined by (9)) will be denoted by $\tilde{F}$.

Now let $\mathfrak{D}(V,F)\ni z\leftrightarrow(V_+,V_-)$. Denoting by V_+^* and V_-^* the images of V_+ and V_- under the mapping $*$, respectively, one gets an orthogonal decomposition of $\tilde{V}$ (with respect to $\tilde{F}$):

$$\tilde{V}=(V_-^*\oplus V_+)+(V_-\oplus V_+^*), \tag{16}$$

where, as is easily seen, $W=V_-^*\oplus V_+$ satisfies the condition (12) with respect to $\tilde{A}$ and $\tilde{F}$. In this way, one obtains a canonical imbedding of $\mathfrak{D}=\mathfrak{D}(V,F)=(I)_{p,q}$ into $\mathfrak{D}'=\mathfrak{D}(\tilde{V}_R,\tilde{A})=(III)_{p+q}$ satisfying clearly the conditions (i), (ii). If $(e_1,\cdots,e_{p+q})$ is an orthonormal basis of V as given in 1. 3, then

$$(e_{p+1}^*,\cdots,e_{p+q}^*,e_1,\cdots,e_p,e_{p+1},\cdots,e_{p+q},e_1^*,\cdots,e_p^*)$$

is an orthonormal basis of $\tilde{V}$ as given in 1. 4. In terms of these basis, the

above imbedding is expressed as $Z \to \begin{pmatrix} 0 & {}^tZ \\ Z & 0 \end{pmatrix}$; and the corresponding monomorphism of Lie algebras, which we shall call $\iota_{p,q}$, is given as follows:

$$(17) \qquad (I)_{p,q} \ni X = \begin{pmatrix} X_1 & X_{12} \\ {}^t\bar{X}_{12} & X_2 \end{pmatrix} \longrightarrow \begin{pmatrix} \bar{X}_2 & 0 & 0 & {}^tX_{12} \\ 0 & X_1 & X_{12} & 0 \\ 0 & {}^t\bar{X}_{12} & X_2 & 0 \\ \bar{X}_{12} & 0 & 0 & \bar{X}_1 \end{pmatrix} \in (III)_{p+q}.$$

It is clear that $\iota_{p,q}$, viewed as a representation of $\mathfrak{g}$ into $M_{p+q}(\boldsymbol{C})$, is equivalent to (id.) $\dotplus$ $(\overline{\text{id.}})$, (id.) denoting the identical representation.

Moreover, note that there is a (non-canonical) analytic isomorphism between $\mathfrak{D} = (I)_{p,q}$ and $\mathfrak{D}' = (I)_{q,p}$ given by the correspondence: $Z \to {}^tZ$; the corresponding isomorphism of Lie algebras is given as follows:

$$(18) \qquad (I)_{p,q} \ni X \longrightarrow J_{p,q}{}^{-1}\bar{X}J_{p,q} \in (I)_{q,p},$$

where

$$J_{p,q} = \begin{pmatrix} 0 & 1_p \\ 1_q & 0 \end{pmatrix}.$$

Then, it can easily be verified that, for $X \in (I)_{p,q}$, we have

$$(18a) \qquad \iota_{q,p}(J_{p,q}{}^{-1}\bar{X}J_{p,q}) = \begin{pmatrix} J_{p,q} & 0 \\ 0 & J_{p,q} \end{pmatrix} \iota_{p,q}(X) \begin{pmatrix} J_{p,q} & 0 \\ 0 & J_{p,q} \end{pmatrix}^{-1}.$$

§ 2. Reduction of the problem.

In this section, we shall fix once and for all the matricial expression of $\mathfrak{g}' = (I)_{p,q}$ or $(III)_p$ as given in 1.3, 4, and regard homomorphisms ρ of $\mathfrak{g}$ into $\mathfrak{g}'$ as matrix representations. We understand by a 'trivial' representation a representation ρ (of any dimension) of $\mathfrak{g}$ such that $\rho(\mathfrak{g}) = \{0\}$. Furthermore, we make the following convention: if ρ_i $(1 \leqq i \leqq r)$ are representations of $\mathfrak{g}$ into $(I)_{p_i,q_i}$ with $\rho_i(X) = \begin{pmatrix} X_1^{(i)} & X_{12}^{(i)} \\ {}^t\bar{X}_{12}^{(i)} & X_2^{(i)} \end{pmatrix}$, the representation of $\mathfrak{g}$ into $(I)_{p,q}$ $(p = \sum p_i,\ q = \sum q_i)$ given by

$$\mathfrak{g} \ni X \longrightarrow \begin{bmatrix} X_1^{(1)} & & & X_{12}^{(1)} & & \\ & \ddots & & & \ddots & \\ & & X_1^{(r)} & & & X_{12}^{(r)} \\ {}^t\bar{X}_{12}^{(1)} & & & X_2^{(1)} & & \\ & \ddots & & & \ddots & \\ & & {}^t\bar{X}_{12}^{(r)} & & & X_2^{(r)} \end{bmatrix}$$

will be called the 'direct sum' of $\rho_1, \cdots, \rho_r$ and denoted by $\rho_1 \dotplus \cdots \dotplus \rho_r$.

If the ρ_i's are representations into $(III)_{p_i}$, then clearly $\rho_1 \dot{+} \cdots \dot{+} \rho_r$ is a representation into $(III)_p$ $(p=\sum p_i)$.

2. 1. Our first step is to reduce the problem to the case where ρ satisfies the stronger condition (H_2).

PROPOSITION 1. *Let $\mathfrak{g}$ be a semi-simple Lie algebra of hermitian type and let ρ be a representation of $\mathfrak{g}$ into $\mathfrak{g}'=(III)_{n'}$ satisfying (H_1). Then there exist a representation ρ_1 of $\mathfrak{g}$ into $(III)_{n_1'}$ $(n_1'\geqq 0)$ satisfying (H_2) and representations ρ_i $(2\leqq i\leqq r)$ of $\mathfrak{g}$ into $(I)_{p_i,q_i}$ satisfying (H_2) with $p_i>0$, $1>q_2/p_2>\cdots>q_r/p_r\geqq 0$ such that we have*

$$\rho \underset{(*)}{\sim} \rho_1 \dot{+} \iota_{p_2,q_2}\circ\rho_2 \dot{+} \cdots \dot{+} \iota_{p_r,q_r}\circ\rho_r \dot{+} (\text{triv.}). \tag{19}$$

Proof. Consider the centralizer $\mathfrak{g}''=\mathfrak{z}(\rho(H_0)-H_0')$. As remarked in 1.2, we may assume that ρ satisfies the condition (4). Then, since we have $\rho(H_0)-H_0'\in\mathfrak{h}'$, $\mathfrak{g}''$ is a reductive subalgebra of $\mathfrak{g}'$. More explicitly, denoting by $\mathfrak{r}''$ the ('closed') subsystem of the root system $\mathfrak{r}'$ of $\mathfrak{g}'_C$ relative to $\mathfrak{h}'_C$ formed of all $\alpha\in\mathfrak{r}'$ with $\alpha(\rho(H_0)-H_0')=0$, we have

$$\mathfrak{g}''_C=\mathfrak{h}'_C+\sum_{\alpha\in\mathfrak{r}''}\mathfrak{g}'_\alpha, \qquad \mathfrak{g}''=\mathfrak{g}''_C\cap\mathfrak{g}',$$

where $\mathfrak{g}'_\alpha$ denotes the eigen-space corresponding to $\alpha\in\mathfrak{r}'$. It follows that the semi-simple part of $\mathfrak{g}''$ is given by $\mathfrak{g}'(\mathfrak{r}'')=\mathfrak{g}'_C(\mathfrak{r}'')\cap\mathfrak{g}'$, $\mathfrak{g}'_C(\mathfrak{r}'')$ denoting the (complex) semi-simple subalgebra (defined over $\boldsymbol{R}$) of $\mathfrak{g}_C$ generated by the $\mathfrak{g}'_\alpha$ $(\alpha\in\mathfrak{r}'')$. Moreover, the Cartan involution τ' of $\mathfrak{g}'$, corresponding to the decomposition $\mathfrak{g}'=\mathfrak{k}'+\mathfrak{p}'$, leaves $\rho(H_0)-H_0'$ invariant, so that τ' leaves also $\mathfrak{g}''=\mathfrak{z}(\rho(H_0)-H_0')$ invariant, thus inducing on $\mathfrak{g}'(\mathfrak{r}'')$ a Cartan involution of it. (Cf. [0], Lemma 1.5. In fact, τ' induces a Cartan involution on each $\boldsymbol{R}$-simple component of $\mathfrak{g}'(\mathfrak{r}'')$.) Let $\mathfrak{g}'(\mathfrak{r}'')=\mathfrak{k}''+\mathfrak{p}''$ be the corresponding Cartan decomposition. If H_0'' denotes the projection of H_0' on the semi-simple part $\mathfrak{g}'(\mathfrak{r}'')$, it is clear that H_0'' is an element in the center of $\mathfrak{k}''$ defining a complex structure on $\mathfrak{p}''$. This implies that $\mathfrak{g}'(\mathfrak{r}'')$ is of hermitian trpe, (and therefore all its $\boldsymbol{R}$-simple components are absolutely simple).

Now to determine the simple components of $\mathfrak{g}'(\mathfrak{r}'')$ more explicitly, put

$$\rho(H_0)=\operatorname{diag.}(\sqrt{-1}\,\lambda_1,\cdots,\sqrt{-1}\,\lambda_{n'},-\sqrt{-1}\,\lambda_1,\cdots,-\sqrt{-1}\,\lambda_{n'})$$

with $\lambda_i\in\boldsymbol{R}$. Then, since $\mathfrak{r}'=\{\pm\xi_i\pm\xi_j, \pm 2\xi_i\ (1\leqq i,j\leqq n', i\neq j)\}$, the decomposition of $\mathfrak{r}''$ is obtained in the following way. Put $N'=\{1,2,\cdots,n'\}$ and

$$N_1'=\{i\in N' \mid \lambda_i=\tfrac{1}{2}\};$$

then decompose $N'-N_1'$ into equivalence-classes with respect to the equivalence-relation defined by

$$i\sim j \Longleftrightarrow \lambda_i-\tfrac{1}{2}=\pm(\lambda_j-\tfrac{1}{2}).$$

Call $N_2',\cdots,N_r'$ the equivalence-classes containing more than one element; then each N_k' $(2\leqq k\leqq r)$ is written in the form $N_k'=P_k\cup Q_k$ with $P_k=\{i\in N_k' \mid \lambda_i<\tfrac{1}{2}\}$ and $Q_k=\{i\in N_k' \mid \lambda_i>\tfrac{1}{2}\}$. We call n_1', p_k, q_k the numbers of elements in N_1', P_k, Q_k, respectively. Then $\mathfrak{r}''$ decomposes into the union of irreducible subsystems as follows:

$$\mathfrak{r}''=\mathfrak{r}''_1\cup\mathfrak{r}''_2\cup\cdots\cup\mathfrak{r}''_r$$

where

$$\begin{aligned}\mathfrak{r}''_1&=\{\pm\xi_i\pm\xi_j\ (i\neq j),\ \pm 2\xi_i \mid i,j\in N_1'\},\\ \mathfrak{r}''_k&=\{\xi_i-\xi_j\ (i\neq j) \mid i,j\in P_k \text{ or } i,j\in Q_k\}\\ &\quad\cup\{\pm(\xi_i+\xi_j) \mid i\in P_k, j\in Q_k\} \qquad (2\leqq k\leqq r).\end{aligned}$$

From what we have proved above, each subalgebra $\mathfrak{g}'_C(\mathfrak{r}''_k)$ is defined over $\boldsymbol{R}$ and $\mathfrak{g}'(\mathfrak{r}''_k)=\mathfrak{g}'_C(\mathfrak{r}''_k)\cap\mathfrak{g}'$ is a Lie algebra corresponding to an irreducible symmetric domain. It follows at once that $\mathfrak{g}'(\mathfrak{r}''_1)$ is $(III)_{n_1'}$ and $\mathfrak{g}'(\mathfrak{r}''_k)$ $(2\leqq k\leqq r)$ are $(I)_{p_k,q_k}$. Moreover, making a permutation of the form

$$\begin{pmatrix} N_1', Q_2, P_2,\cdots, Q_r, P_r, *** \\ 1, 2, \cdots\cdots\cdots\cdots, n' \end{pmatrix},$$

(N_1', P_k and Q_k being ordered arbitrarily), which is clearly effected by an inner automorphism defined by an element in K', we have

$$\mathfrak{g}'(\mathfrak{r}'')=(III)_{n_1'}\dotplus\iota_{p_2,q_2}((I)_{p_2,q_2})\dotplus\cdots\dotplus\iota_{p_r,q_r}((I)_{p_r,q_r})\dotplus(0).$$

Now the condition (H_1) is equivalent to saying that $\rho(\mathfrak{g})\subset\mathfrak{z}(\rho(H_0)-H_0')=\mathfrak{g}''$; since $\rho(\mathfrak{g})$ is semi-simple, it is then contained in $\mathfrak{g}'(\mathfrak{r}'')$, (which implies that $H_0''=\rho(H_0)$). Thus, denoting by $\rho_k(X)$ $(X\in\mathfrak{g})$ the projection of $\rho(X)\in\mathfrak{g}'(\mathfrak{r}'')$ on $\mathfrak{g}'(\mathfrak{r}''_k)$, we obtain representations ρ_k of $\mathfrak{g}$ into $(III)_{n_1'}$ or $(I)_{p_k,q_k}$ $(2\leqq k\leqq r)$ satisfying (19). Since $\rho_k(H_0)$ coincide with the projections of $H_0''=\rho(H_0)$ on $\mathfrak{g}'(\mathfrak{r}''_k)$, the ρ_k's satisfy the condition (H_2). Finally, we notice that $\lambda_i=\dfrac{q_k}{p_k+q_k}$ for $i\in P_k$ and $=\dfrac{p_k}{p_k+q_k}$ for $i\in Q_k$; therefore, we may assume that $1>q_2/p_2>\cdots>q_r/p_r$, q.e.d.

2. 2. Our next task is to reduce the problem to the case of absolutely irreducible representation. $\mathfrak{g}$ being, as before, a semi-simple Lie algebra of hermitian type, suppose first that a representation ρ of $\mathfrak{g}$ into $\mathfrak{g}'=(I)_{p,q}$

$=\mathfrak{g}(V,F)$ satisfying (H_2) is given, where (V,F) is the hermitian vector-space on which $\mathfrak{g}'$ operates. (We include the definite case where p or $q=0$.) We shall show that (V,F) is completely reducible (as hermitian vector-space) under $\rho(\mathfrak{g})$, i.e. V is an orthogonal sum of minimal $\rho(\mathfrak{g})$-invariant subspaces.

Let V_1 be a minimal $\rho(\mathfrak{g})$-invariant subspace ($\neq\{0\}$) of V. As the hermitian form F is $\rho(\mathfrak{g})$-invariant (i.e. one has $F(\rho(X)x,y)+F(x,\rho(X)y)=0$ for all $X\in\mathfrak{g}$, $x,y\in V$), the orthogonal complement $V_1^{\perp}$ of V_1 is also $\rho(\mathfrak{g})$-invariant. As V_1 is minimal, this implies that $V_1\cap V_1^{\perp}$ is either $=V_1$ or $=0$; in the second case, V is the direct sum of V_1 and $V_1^{\perp}$. So, in order to prove our assertion, it is enough to show that the first case does not occur, or, in other words, that $F_1=F\mid V_1$ is not identically zero. In the notation of 1.3, we have

$$F_u(x,y)=F(Tx,y) \qquad \text{for all } x,y\in V,$$

where T is given as follows (Lem. 1):

$$T=-2iH_0'+\frac{p-q}{p+q}1_V. \tag{20}$$

V_1 being invariant under $\rho(H_0)=H_0'$, this implies that V_1 is also invariant under T. Therefore, if $F_1=0$, we would have $F_{u1}=F_u\mid V_1=0$, which is impossible, for F_u is a definite hermitian form. This proves our assertion.

One sees, at the same time, that F_1 and F_{u1} stand in the same relation as F and F_u, by means of the orthogonal decomposition: $V_1=(V_1\cap V_+^{(0)})+(V_1\cap V_-^{(0)})$. Therefore, if F_1 has signature (p_1,q_1) and if $H_0{}'_1$ denotes the element in the corresponding maximal compact subalgebra of $\mathfrak{g}(V_1,F_1)$, defining the complex structure, one has

$$T\mid V_1=-2iH_0{}'_1+\frac{p_1-q_1}{p_1+q_1}1_{V_1}.$$

Comparing this with the relation obtained from (20) by restricting it on V_1, and noting that $\mathrm{tr}(H_0'\mid V_1)=\mathrm{tr}(H_0{}'_1)=0$, one concludes that $H_0{}'_1=H_0'\mid V_1$ and $p_1:q_1=p:q$. It follows that $\rho(H_0)\mid V_1=H_{01}'$, i.e. $\rho_1=\rho\mid V_1$ satisfies the condition (H_2).

Translating these results into the language of representation-theory, we obtain the following

PROPOSITION 2. *Let $\mathfrak{g}$ be a semi-simple Lie algebra of hermitian type and let ρ be a representation of $\mathfrak{g}$ into $\mathfrak{g}'=(I)_{p,q}$ satisfying (H_2). Then there exist non-negative integers p_i, q_i $(1\leqq i\leqq s)$ with $p_i:q_i=p:q$,*

$\sum p_i = p$, $\sum q_i = q$ and absolutely irreducible representations ρ_i of $\mathfrak{g}$ into $(I)_{p_i, q_i}$ satisfying (H_2) such that ρ is (k)-equivalent to the direct sum of the ρ_i $(1 \leqq i \leqq s)$.

2.3. The notations $\mathfrak{g}$, ρ, V, $F, \cdots$ being as before, we next consider the case where $\rho(\mathfrak{g})$ is contained in $(III)_p$. In this case, we shall show that, in Prop. 2 (with $p = q$), we may choose the representations ρ_i to be either

a) an absolutely irreducible representation of $\mathfrak{g}$ into $(III)_{p_i}$, or

b) an absolutely irreducible representation of $\mathfrak{g}$ into $(I)_{p_i, p_i}$ appearing in a pair $(\rho_i, \bar{\rho}_i)$, such that $\iota_{p_i, p_i} \circ \rho_i \sim \rho_i \dotplus \bar{\rho}_i$ is an $\boldsymbol{R}$-irreducible representation of g into $(III)_{2p_i}$.

To prove this, let $V_{\boldsymbol{R}}$ be the real form of V and A the non-degenerate alternating form on V satisfying (8), defining $(III)_p \subset (I)_{p,p}$. Let V_1 be a minimal $\rho(\mathfrak{g})$-invariant subspace of V. By the same reason as in 2.2, $A \mid V_1$ is either non-degenerate or identically zero. Suppose first that $A \mid V_1$ is non-degenerate. If $\bar{V}_1 = V_1$, we have clearly the case a). If not, we shall show that V_1 can be replaced by an isomorphic subspace satisfying this condition. In fact, for each $x \in V_1$, there exists uniquely $x' \in V_1$ such that we have

$$A(\bar{x}, y) = A(x', y) \qquad \text{for all } y \in V_1;$$

and the correspondence $\phi\colon x \to x'$ is a semi-linear transformation of V_1 into itself, which clearly commutes with $\rho(X)$ $(X \in \mathfrak{g})$. Hence, by Schur's lemma, ϕ^2 is a scalar $\lambda 1_{V_1}$. Here we have $\lambda > 0$; for, since $A(\bar{V}_1, V_1) = F(V_1, V_1) \neq 0$, we have $\lambda \neq 0$, and, in view of the relations $F_u(x, y) = F(Tx, y) = iA(\overline{Tx}, y)$ and $\bar{T} = -T$, we have, for any $x \in V_1$,

$$\begin{aligned} F_u(\phi(x), \phi(x)) &= -iA(T\phi^2(x), \phi(x)) = -\lambda iA(Tx, \phi(x)) \\ &= -\lambda iA(Tx, \bar{x}) = \lambda F_u(x, x), \end{aligned}$$

whence follows $\lambda > 0$. Therefore, if we put $x^{\sigma_1} = \lambda^{-\frac{1}{2}}\phi(x)$ for $x \in V_1$, σ_1 is a conjugation of V_1, which is different from the bar. Put further

$$\psi(x) = x - \bar{x}^{\sigma_1} \qquad \text{for } x \in V_1,$$

and $V_1' = \psi(V_1)$; then ψ is a $\boldsymbol{C}$-linear isomorphism of V_1 onto V_1', commuting with $\rho(X)$ $(X \in \mathfrak{g})$ and such that $\psi(x^{\sigma_1}) = -\overline{\psi(x)}$. Thus V_1' is isomorphic with V_1 (as $\rho(\mathfrak{g})$-space) and has the property $\overline{V_1'} = V_1'$. This proves our assertion.

Next, suppose that $A \mid V_1 = 0$. Then, again by $A(\bar{V}_1, V_1) = F(V_1, V_1)$

$\neq 0$, we have $V \cap \bar{V}_1 = 0$; and $\bar{V}_1$ may be identified with the dual space of V_1 by means of the inner product

$$\langle x, y \rangle = A(x, y) \qquad \text{for } x \in \bar{V}_1, y \in V_1.$$

Then, in the notation of 1.5 applied to V_1, F_1, $V_1^* = \bar{V}_1$, we see that σ and $\tilde{A}$ coincide with the restrictions on $V_1 + \bar{V}_1$ of the bar and A, respectively. In other words, the symplectic space $V_1 + \bar{V}_1$ is nothing else than the one constructed from (V_1, F_1) in the manner described in 1.5. Therefore, in a suitable basis of $V_1 + \bar{V}_1$, we have $\rho \mid (V_1 + \bar{V}_1) = \iota_{p_1, p_1} \circ \rho_1$. Here we may suppose that $V_1 + \bar{V}_1$ does not contain any minimal $\rho(\mathfrak{g})$-invariant subspace V_1' such that $\bar{V}_1' = V_1'$, for otherwise this case is again reduced to the previous case a). Then we are in the case b).

Thus we have proved that V can be decomposed into the orthogonal sum (with respect to F as well as to A) of minimal $\rho(\mathfrak{g})$-invariant subspaces as follows:

$$V = \sum_{i=1}^{s_1} V_i + \sum_{i=s_1+1}^{s_1+s_2} (V_i + \bar{V}_i),$$

where V_i $(1 \leqq i \leqq s_1)$ are defined over $\boldsymbol{R}$, i.e. such that $\bar{V}_i = V_i$, and $V_i + \bar{V}_i$ $(s_1 + 1 \leqq i \leqq s_1 + s_2)$ contain no proper $\rho(\mathfrak{g})$-invariant subspace defined over $\boldsymbol{R}$.

This result can also be formulated as follows:

PROPOSITION 3. *Let $\mathfrak{g}$ be a semi-simple Lie algebra of hermitian type and let ρ be a representation of $\mathfrak{g}$ into $\mathfrak{g}' = (III)_p$ satisfying (H_2). Then there exist positive integers p_i $(1 \leqq i \leqq s_1 + s_2)$ with $\sum_{i=1}^{s_1} p_i + 2 \sum_{i=s_1+1}^{s_1+s_2} p_i = p$, absolutely irreducible representations ρ_i $(1 \leqq i \leqq s_1)$ of $\mathfrak{g}$ into $(III)_{p_i}$ satisfying (H_2) and absolutely irreducible representations ρ_i $(s_1 + 1 \leqq i \leqq s_1 + s_2)$ of $\mathfrak{g}$ into $(I)_{p_i, p_i}$ satisfying (H_2), which are not (k)-equivalent (in $(I)_{p_i, p_i}$) to representations contained in $(III)_{p_i}$, such that ρ is (k)-equivalent (in $(III)_p$) to the direct sum of the ρ_i $(1 \leqq i \leqq s_i)$ and the $\iota_{p_i, p_i} \circ \rho_i$ $(s_1 + 1 \leqq i \leqq s_1 + s_2)$.*

2.4. Combining Prop. 1, 2 and 3, we obtain the first part of the following theorem; the third ('converse') part of it is trivial.

THEOREM 1. *Let $\mathfrak{g}$ be a semi-simple Lie algebra of hermitian type and let ρ be a representation of $\mathfrak{g}$ into $(III)_{n'}$ satisfying the condition (H_1). Then there exist absolutely irreducible representations ρ_i $(1 \leqq i \leqq r_1)$ of $\mathfrak{g}$ into $(III)_{p_i}$ $(p_i > 0)$ satisfying (H_2) and absolutely irreducible representations ρ_i $(r_1 + 1 \leqq i \leqq r_1 + r_2)$ of $\mathfrak{g}$ into $(I)_{p_i, q_i}$ $(p_i, q_i \geqq 0, p_i + q_i > 0)$ satisfying*

(H_2) such that ρ is (k)-equivalent to the direct sum of the ρ_i $(1 \leqq i \leqq r_1)$, the $\iota_{p_i,q_i} \circ \rho_i$ $(r_1+1 \leqq i \leqq r_1+r_2)$ and a trivial representation:

$$\rho \underset{(k)}{\sim} \sum_{i=1}^{r_1} \rho_i \dotplus \sum_{i=r_1+1}^{r_1+r_2} \iota_{p_i,q_i} \circ \rho_i \dotplus (\text{triv.}). \tag{21}$$

If, moreover, we assume that the ρ_i $(r_1+1 \leqq i \leqq r_1+r_2)$ with $p_i = q_i$ are not (k)-equivalent to representations contained in $(III)_{p_i}$, then the above decomposition is unique up to the order, (k)-equivalence (in each $(III)_{p_i}$ or $(I)_{p_i,q_i}$) and the replacement of ρ_i by $J_{p_i,q_i}{}^{-1}\bar{\rho}_i J_{p_i,q_i}$ for $r_1+1 \leqq i \leqq r_1+r_2$ where $J_{p_i,q_i} = \begin{pmatrix} 0 & 1_{p_i} \\ 1_{q_i} & 0 \end{pmatrix}$. Conversely, any representation ρ given in this form is a representation of $\mathfrak{g}$ into $(III)_{n'}$ satisfying (H_1).

To obtain the uniqueness, we first prove the following

LEMMA 2. *Let $\mathfrak{g}$ be a semi-simple Lie algebra of hermitian type and let ρ_1 and ρ_2 be absolutely irreducible representations of $\mathfrak{g}$ into $(I)_{p,q}$ and $(I)_{p',q'}$, respectively, satisfying (H_2). If $p+q=p'+q'$ and ρ_1, ρ_2 are equivalent (in the usual sense), then we have $p=p'$, $q=q'$, and ρ_1, ρ_2 are (k)-equivalent in $(I)_{p,q}$. If, moreover, $p=q$ and the $\rho_i(\mathfrak{g})$'s are both contained in $(III)_p$ (or $(II)_p$), then they are (k)-equivalent in $(III)_p$ (or $(II)_p$, resp.).*

Proof. By the assumption, we have $\rho_2(X) = P^{-1}\rho_1(X)P$ for all $X \in \mathfrak{g}$ with a non-singular matrix P of degree $p+q$. Then, by the condition (H_2), we have $P^{-1}H_0'P = H_0''$ where

$$H_0' = \operatorname{diag.}\left(\frac{q}{p+q}\, i1_p, \quad -\frac{p}{p+q}\, i1_q\right),$$

$$H_0'' = \operatorname{diag.}\left(\frac{q'}{p'+q'}\, i1_{p'}, \quad -\frac{p'}{p'+q'}\, i1_{q'}\right),$$

whence follows that $p=p'$, $q=q'$ and that P is of the form $\operatorname{diag.}(P_1, P_2)$ with $P_1 \in GL(p, \boldsymbol{C})$ and $P_2 \in GL(q, \boldsymbol{C})$. Moreover, since $\rho_2(X)$ $(X \in \mathfrak{g})$ leaves the hermitian forms with the matrices $1_{p,q}$ and ${}^t\bar{P}1_{p,q}P$ invariant, it follows by Schur's lemma that ${}^t\bar{P}1_{p,q}P = \lambda 1_{p,q}$ with $\lambda > 0$. Hence, replacing P by $\lambda^{-\frac{1}{2}}P$, we may assume that $P_1 \in U(p)$, $P_2 \in U(q)$, or, in other words, that P (modulo scalar matrices) belongs to the (fixed) maximal compact subgroup of $PU(p,q)$. This proves our first assertion. Now, if $p=q$ and if the ρ_i's are both contained in $(III)_p$, it follows again by Schur's lemma that we have

$${}^tP \begin{pmatrix} 0 & -1_p \\ 1_p & 0 \end{pmatrix} P = \epsilon \begin{pmatrix} 0 & -1_p \\ 1_p & 0 \end{pmatrix},$$

whence follows that ${}^tP_2P_1=\epsilon 1_p$. Since we have $\epsilon\bar{\epsilon}=1$, we can write $\epsilon=\eta^2$ with $\eta\bar{\eta}=1$; hence, replacing P by $\eta^{-1}P$, we may assume that $P_2={}^tP_1^{-1}=P_1$. This proves that ρ_1 and ρ_2 are (k)-equivalent in $(III)_p$. The case where the ρ_i's are contained in $(II)_p$ can be treated quite similarly, q. e. d.

LEMMA 3. *Let* $\mathfrak{g}$ *be a semi-simple Lie algebra of hermitian type,* ρ *a representation of* $\mathfrak{g}$ *into* $(I)_{p,q}$ *satisfying* (H_2) *and let* $J_{p,q}=\begin{pmatrix} 0 & 1_p \\ 1_q & 0 \end{pmatrix}$. *Then* $J_{p,q}^{-1}\bar{\rho}J_{p,q}$ *is a representation of* $\mathfrak{g}$ *into* $(I)_{q,p}$ *satisfying* (H_2) *and we have*

$$\iota_{p,q}\circ\rho \underset{(k)}{\sim} \iota_{q,p}\circ(J_{p,q}^{-1}\bar{\rho}J_{p,q}) \qquad (in\ (III)_{p+q}).$$

This is an immediate consequence of what we have stated at the end of 1. 5.

We shall now show the uniqueness of the decomposition (21). Suppose that we have another decomposition

$$\rho\sim\sum_{i=1}^{r_1'}\rho_i' \dotplus \sum_{i=r_1'+1}^{r_1'+r_2'}\iota_{p_i',q_i'}\circ\rho_i' \dotplus (\text{triv.})$$

with the properties described in Th. 1. Then it is clear that $r_1+2r_2=r_1'+2r_2'$ and that there is a one-to-one correspondence between irreducible factors

$$\rho_i\ (1\leqq i\leqq r_1+r_2),\ \bar{\rho}_i\ (r_1+1\leqq i\leqq r_1+r_2)$$
$$\text{and } \rho_i'\ (1\leqq i\leqq r_1'+r_2'),\ \bar{\rho}_i'\ (r_1'+1\leqq i\leqq r_1'+r_2')$$

such that the corresponding representations are equivalent. Now suppose that ρ_i is corresponding to $\rho_{i'}'$ or $\bar{\rho}_{i'}'$. Then it follows from Lem. 2 that we have $1\leqq i\leqq r_1$ if and only if $1\leqq i'\leqq r_1'$ and that, if that is so, ρ_i and $\rho_{i'}'$ are (k)-equivalent in $(III)_{p_i}$ $(p_i=p_{i'}')$. Thus $r_1=r_1'$, $r_2=r_2'$ and we may suppose that $i=i'$. Now let $r_1+1\leqq i\leqq r_1+r_2$. If ρ_i is equivalent to ρ_i', we have (by Lem. 2) $p_i=p_i'$, $q_i=q_i'$, and ρ_i and ρ_i' are (k)-equivalent in $(I)_{p_i,q_i}$. If ρ_i is equivalent to $\bar{\rho}_i'$, we have (Lem. 2, 3) $p_i=q_i'$, $q_i=p_i'$, and ρ_i is (k)-equivalent to $J_{q_i,p_i}^{-1}\rho_i'J_{q_i,p_i}$ in $(I)_{p_i,q_i}$. But then, by virtue of Lem. 3, replacing ρ_i' by $J_{q_i,p_i}^{-1}\bar{\rho}_i'J_{q_i,p_i}$, we may reduce this case to the previous one. Th. 1 is thereby proved completely.

From the above considerations, we also obtain the following

COROLLARY. *Let* $\mathfrak{g}$ *be a semi-simple Lie algebra of hermitian type. Then two representations* ρ_1 *and* ρ_2 *of* $\mathfrak{g}$ *into* $(III)_{n'}$ *satisfying* (H_1) *are equivalent (in the usual sense), if and only if they are* (k)*-equivalent in* $(III)_{n'}$.

2. 5. Finally let us reduce the problem to the case where $\mathfrak{g}$ is simple. Suppose that $\mathfrak{g}$ is not simple and decomposable as $\mathfrak{g}=\mathfrak{g}_1\oplus\mathfrak{g}_2$, $\mathfrak{g}_1$ and $\mathfrak{g}_2$

being ideals of $\mathfrak{g}$. Let ρ be an absolutely irreducible representation of $\mathfrak{g}$ into $\mathfrak{g}'=\mathfrak{g}(V,F)=(I)_{p,q}$ satisfying (H_2). Then, it is well-known that V is a tensor-product of two (complex) vector-spaces V_i $(i=1,2)$, giving absolutely irreducible representations ρ_i of $\mathfrak{g}_i$, respectively, such that we have

$$(22) \qquad \rho(X_1+X_2)=\rho_1(X_1)\otimes 1_{V_2}+1_{V_1}\otimes\rho_2(X_2) \qquad \text{for } X_i\in\mathfrak{g}_i,$$

1_{V_i} denoting the identity transformation of V_i.

In the first place, we shall show that there exist $\rho_i(\mathfrak{g}_i)$-invariant (non-degenerate) hermitian forms F_i $(i=1,2)$ on V_i such that we have

$$(23) \qquad F(x_1\otimes x_2, y_1\otimes y_2)=F_1(x_1,y_1)\cdot F_2(x_2,y_2) \qquad \text{for } x_i, y_i\in V_i.$$

In fact, F can be expressed as a finite sum

$$F=\sum_\alpha F_1{}^{(\alpha)}\otimes F_2{}^{(\alpha)}$$

of sesqui-linear forms $F_i{}^{(\alpha)}$ on $V_i\times V_i$; we may suppose that the $F_1{}^{(\alpha)}$'s and the $F_2{}^{(\alpha)}$'s are both linearly independent over $\boldsymbol{C}$. Then from the invariance of F under $\rho(\mathfrak{g})$ follows the invariance of $F_i{}^{(\alpha)}$ under $\rho_i(\mathfrak{g}_i)$. ρ_2 being absolutely irreducible, this implies by Schur's lemma that the $F_2{}^{(\alpha)}$'s are all non-degenerate and are scalar multiples of one of them. Thus we may write $F=F_1\otimes F_2$ with $\rho_i(\mathfrak{g}_i)$-invariant sesqui-linear forms F_i on $V_i\times V_i$. Now, since F is hermitian, we have the relations:

$$F_1(y_1,x_1)=\epsilon\overline{F(x_1,y_1)}, \qquad F_2(y_2,x_2)=\epsilon^{-1}\overline{F_2(x_2,y_2)}$$

with a complex number ϵ with $\epsilon\bar{\epsilon}=1$. Writing ϵ in the form $\epsilon=\eta\bar{\eta}^{-1}$ and replacing F_1 and F_2 by $\eta^{-1}F_1$ and ηF_2, respectively, we may assume the F_i's to be hermitian. This proves our assertion. It should be noted that F_1 and F_2 are determined uniquely up to real scalar multiples.

In the second place, from the facts that the Cartan involution τ of $\mathfrak{g}$ leaves the ideals $\mathfrak{g}_1$ and $\mathfrak{g}_2$ invariant and that it satisfies the relation $\rho(X^\tau)=\rho(X)^{\tau'}=T^{-1}\rho(X)T$ $(X\in\mathfrak{g})$, one can show by the similar arguments as above that T is a tensor-product of linear transformations T_i $(i=1,2)$ on V_i:

$$(24) \qquad T=T_1\otimes T_2.$$

Then we have

$$F_u(x,y)=F(Tx,y)=F_1(T_1x_1,y_1)\cdot F_2(T_2x_2,y_2)$$

for $x=x_1\otimes x_2$, $y=y_1\otimes y_2$; and, as above, we can adjust T_i's in such a way that $F_{ui}(x_i,y_i)=F_i(T_ix_i,y_i)$ becomes hermitian. Then F_{u1} and F_{u2} are clearly definite and so, replacing T_i by $-T_i$ if necessary, we may suppose them to be positive-definite. From the relation $1=T^2=T_1{}^2\otimes T_2{}^2$, it follows that $T_1{}^2$ is a scalar $\lambda 1_{V_1}$, where λ is positive, as is seen from the relation

$F_{*1}(T_1x_1, T_1x_1) = \lambda F_{*1}(x_1, x_1)$. Thus, replacing T_1 and T_2 by $\lambda^{-\frac{1}{2}}T_1$ and $\lambda^{\frac{1}{2}}T_2$, respectively, we have $T_i^2 = 1_{V_i}$. T_1 and T_2 satisfying these conditions are uniquely determined.

We have thus proved, firstly, that the hermitian vector-space (V, F) is the tensor-product of two hermitian vector-spaces (V_1, F_1) and (V_2, F_2), and that, if we put $\mathfrak{g}_i' = \mathfrak{g}(V_i, F_i)$ $(i = 1, 2)$, the representation ρ is factorized as follows:

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \xrightarrow{\rho_1 \oplus \rho_2} \mathfrak{g}_1' \oplus \mathfrak{g}_2' \xrightarrow{\kappa} \mathfrak{g}',$$

where κ denotes the canonical injection defined by

$$\kappa: \quad \mathfrak{g}_1' \oplus \mathfrak{g}_2' \ni (X_1', X_2') \to X_1' \otimes 1_{V_2} + 1_{V_1} \otimes X_2' \in \mathfrak{g}'.$$

Secondly, we have shown that, to the orthogonal decomposition $V = V_+^{(0)} + V_-^{(0)}$ (determining F_*), there correspond uniquely orthogonal decompositions $V_i = V_{i+} + V_{i-}$ (determining F_{*i}, respectively) such that we have

$$V_+^{(0)} = V_{1+} \otimes V_{2+} + V_{1-} \otimes V_{2-}, \qquad V_-^{(0)} = V_{1+} \otimes V_{2-} + V_{1-} \otimes V_{2+}.$$

It follows that, if F_i is of signature (p_i, q_i), we have $p = p_1p_2 + q_1q_2$, $q = p_1q_2 + q_1p_2$; and that, if $\mathfrak{g}_i = \mathfrak{k}_i + \mathfrak{p}_i$ is the Cartan decomposition of $\mathfrak{g}_i$ defined by $\tau \mid \mathfrak{g}_i$ and if $\mathfrak{g}_i' = \mathfrak{k}_i' + \mathfrak{p}_i'$ is that of $\mathfrak{g}_i'$ corresponding to T_i, then we have

$$\rho_i(\mathfrak{k}_i) \subset \mathfrak{k}_i', \qquad \rho_i(\mathfrak{p}_i) \subset \mathfrak{p}_i' \qquad\qquad (i = 1, 2).$$

Now let H_{0i} be the projection of H_0 on $\mathfrak{g}_i$ and let $H_0{}'_i$ be the element in the center of $\mathfrak{k}_i'$ defining the complex structure of $\mathfrak{p}_i'$ $(i = 1, 2)$. We shall show that, under the assumption (H_2), one of the F_i's, say F_1, is definite and we have

$$\rho_1(H_{01}) = H_0{}'_1 = 0, \qquad \rho_2(H_{02}) = \pm H_0{}'_2.$$

In fact, from Lem. 1, we have

$$T_i = -2\sqrt{-1}\, H_0{}'_i + \frac{p_i - q_i}{p_i + q_i} 1_{V_i} \qquad\qquad (i = 1, 2)$$

and so, from (24),

$$\begin{aligned}(25) \quad T &= T_1 \otimes T_2 \\ &= -4H_0{}'_1 \otimes H_0{}'_2 - 2\sqrt{-1}\left(\frac{p_2 - q_2}{p_2 + q_2} H_0{}'_1 \otimes 1_{V_2} + \frac{p_1 - q_1}{p_1 + q_1} 1_{V_1} \otimes H_0{}'_2\right) \\ &\quad + \frac{(p_1 - q_1)(p_2 - q_2)}{(p_1 + q_1)(p_2 + q_2)} 1_V.\end{aligned}$$

On the other hand, using the condition (H_2), we have

$$(25') \qquad \begin{aligned} T &= -2\sqrt{-1}H_0' + \frac{p-q}{p+q} 1_V \\ &= -2\sqrt{-1}\,(\rho_1(H_{01}) \otimes 1_{V_2} + 1_{V_1} \otimes \rho_2(H_{02})) + \frac{p-q}{p+q} 1_V. \end{aligned}$$

Since $\mathrm{tr}(\rho_i(H_{0i})) = \mathrm{tr}(H_0{'}_i) = 0$ and since 1_{V_i} and a (non-zero) linear transformation of V_i with trace zero are linearly independent, it follows, by comparison of (25) and (25′), that $H_0{'}_1 \otimes H_0{'}_2 = 0$, which implies $H_0{'}_1$ or $H_0{'}_2 = 0$. Hence, assuming $H_0{'}_1 = 0$, i.e. p_1 or $q_1 = 0$, we obtain further $\rho_1(H_{01}) = 0$ and $\rho_2(H_{02}) = \pm H_{02}'$ (according as $q_1 = 0$ or $p_1 = 0$), which proves our assertion. Replacing F_i by $-F_i$ if necessary, we may assume that F_1 is positive-definite. Then we have $\rho_2(H_{02}) = H_0{'}_2$, i.e. ρ_2 satisfies the condition (H_2).

From these considerations and by an easy induction, we can conclude the following

THEOREM 2. *Let* $\mathfrak{g}$ *be a semi-simple Lie algebra of hermitian type and let*

$$(26) \qquad \mathfrak{g} = \mathfrak{g}_0 + \sum_{i=1}^{m} \mathfrak{g}_i$$

be a decomposition of $\mathfrak{g}$ *into the direct sum of ideals, where* $\mathfrak{g}_0$ *is compact and* $\mathfrak{g}_i$ $(1 \leqq i \leqq m)$ *are simple and non-compact. Let* ρ *be an absolutely irreducible representation of* $\mathfrak{g}$ *into* $(I)_{p,q}$ *satisfying* (H_2). *Then there exist an absolutely irreducible representation* ρ_0 *of* $\mathfrak{g}_0$ *into* $(I)_{n_0,0}$ $(n_0 > 0)$ *and, for some* i_0 $(1 \leqq i_0 \leqq m)$, *an absolutely irreducible representation* ρ_{i_0} *of* $\mathfrak{g}_{i_0}$ *into* $(I)_{p_0,q_0}$ *with* $p = n_0 p_0$, $q = n_0 q_0$ *satisfying* (H_2) *such that* ρ *is* (k)-*equivalent to the representation of the following form*:

$$(27) \qquad \rho(\sum_{i=0}^{m} X_i) = \rho_0(X_0) \otimes 1_{p_0+q_0} + 1_{n_0} \otimes \rho_{i_0}(X_{i_0}) \ \textit{for}\ X_i \in \mathfrak{g}_i \ (0 \leqq i \leqq m).$$

Conversely, any representation of this form is an absolutely irreducible representation of $\mathfrak{g}$ *into* $(I)_{p,q}$ *satisfying* (H_2).

In (27), $Y_0 \otimes \begin{array}{cc} \begin{array}{cc} \overbrace{\quad}^{p_0} & \overbrace{\quad}^{q_0} \end{array} & \\ \begin{pmatrix} Y_1 & Y_{12} \\ {}^t\bar{Y}_{12} & Y_2 \end{pmatrix} & \begin{array}{l} \}\,p_0 \\ \}\,q_0 \end{array} \end{array}$ means $\begin{pmatrix} Y_0 \otimes Y_1 & Y_0 \otimes Y_{12} \\ Y_0 \otimes {}^t\bar{Y}_{12} & Y_0 \otimes Y_2 \end{pmatrix}$.

2. 6. We shall now examine the condition under which the representation ρ in Th. 2 (with $p = q$) is contained in $(III)_p$. The notations being as in 2. 5, let $x \to \bar{x}$ and A be the conjugation and the (non-degenerate)

alternating form on V, respectively, defining $\mathfrak{g}(V_R, A) = (III)_p$, and suppose that $\rho(\mathfrak{g})$ is contained in $\mathfrak{g}(V_R, A)$. Then, by similar arguments as in 2.5, one can prove that there exist $\rho_i(\mathfrak{g}_i)$-invariant (non-degenerate) bilinear forms B_i $(i=1,2)$ on $V_i \times V_i$ such that we have

$$(28) \qquad A(x_1 \otimes x_2, y_1 \otimes y_2) = B_1(x_1, y_1) \cdot B_2(x_2, y_2),$$

where one of B_i's is symmetric and the other is alternating. (B_i's are unique up to (complex) scalar multiple.) Similarly, there exist (non-singular) semi-linear transformations σ_i $(i=1,2)$ on V_i commuting with $\rho_i(\mathfrak{g}_i)$ and such that we have

$$(29) \qquad \overline{x_1 \otimes x_2} = x_1^{\sigma_1} \otimes x_2^{\sigma_2}.$$

It follows that σ_i^2 is a real scalar $\lambda_i 1_{V_i}$ with $\lambda_1\lambda_2 = 1$; hence, replacing σ_i by $|\lambda_i|^{-\frac{1}{2}}\sigma_i$, we may suppose that $\lambda_1 = \lambda_2 = \pm 1$. On the other hand, from the relation $\overline{A(x_1 \otimes x_2, y_1 \otimes y_2)} = A(\overline{x_1 \otimes x_2}, \overline{y_1 \otimes y_2})$, we have $B_i(x_i^{\sigma_i}, y_i^{\sigma_i}) = \epsilon_i \overline{B_i(x_i, y_i)}$ with $\epsilon_1\epsilon_2 = 1$, and, using the relation $\sigma_i^2 = \pm 1_{V_i}$, we obtain $\epsilon_i\bar{\epsilon}_i = 1$. Hence, writing $\lambda_i^{-1}\epsilon_i$ in the form $\lambda_i^{-1}\epsilon_i = \eta_i\bar{\eta}_i^{-1}$ with $\eta_1\eta_2 = 1$ and replacing B_i by $\eta_i^{-1}B_i$, we may suppose that

$$(30) \qquad B_i(x_i^{\sigma_i}, y_i^{\sigma_i}) = \lambda_i \overline{B_i(x_i, y_i)}.$$

Then, we see that, according as B_i is symmetric or alternating, $B_i(x_i^{\sigma_i}, y_i)$ or $\sqrt{-1}\,B_i(x_i^{\sigma_i}, y_i)$ is hermitian. Therefore, in view of the relation

$$F(x_1 \otimes x_2, y_1 \otimes y_2) = F_1(x_1, y_1) \cdot F_2(x_2, y_2) = \sqrt{-1}\,A(\overline{x_1 \otimes x_2}, y_1 \otimes y_2),$$

we conclude that we may suppose that

$$(31) \qquad F_i(x_i, y_i) = \begin{cases} B_i(x_i^{\sigma_i}, y_i) & \text{if } B_i \text{ is symmetric,} \\ \sqrt{-1}\,B_i(x_i^{\sigma_i}, y_i) & \text{if } B_i \text{ is alternating.} \end{cases}$$

Finally, we may suppose, as was shown in 2.5, that F_1 is positive-definite.

Now we have to distinguish two cases:

1°) *B_1 is symmetric and B_2 is alternating.* From the relation

$$F_1(x_1^{\sigma_1}, x_1^{\sigma_1}) = B_1(x_1^{\sigma_1^2}, x_1^{\sigma_1}) = \lambda_1 F_1(x_1, x_1),$$

we infer that $\lambda_1 = 1$. It follows that the σ_i's are conjugations of V_i and that the B_i's are bilinear forms defined over $\boldsymbol{R}$. Therefore ρ_1 is an absolutely irreducible representation into the Lie algebra of the special orthogonal group $SO(V_{1R}, B_1)$ (which is compact) and ρ_2 is an absolutely irreducible representation into $\mathfrak{g}(V_{2R}, B_2) = (III)_p$ satisfying (H_2).

2^0) *B_1 is alternating and B_2 is symmetric.* From the relation

$$F_1(x_1^{\sigma_1}, x_1^{\sigma_1}) = \sqrt{-1}\, B_1(x_1^{\sigma_1^2}, x_1^{\sigma_1}) = -\lambda_1 F_1(x_1, x_1),$$

we infer that $\lambda_1 = -1$. In this case, we can define on V_i a structure of (right) vector-space over the real quaternion-algebra $\boldsymbol{K}$. Namely, in the usual notation, write $\boldsymbol{K} = \boldsymbol{C} + j\boldsymbol{C}$ with $j^2 = -1$, $ij = -ji$, and define the operation of j on V_i by

$$xj = x^{\sigma_i} \qquad \text{for } x \in V_i.$$

Then it is easy to see that V_i's become (right) vector-spaces over $\boldsymbol{K}$. Moreover, if one puts

$$\hat{F}_i(x, y) = \sqrt{-1}(B_i(x^{\sigma_i}, y) + jB_i(x, y)) \qquad \text{for } x, y \in V_i,$$

then one can verify at once that $\hat{F}_1$ (resp. $\hat{F}_2$) becomes a $\boldsymbol{K}$-valued hermitian (resp. anti-hermitian) sesqui-linear form on V_i with respect to the canonical involution of $\boldsymbol{K}$, and that $\hat{F}_1$ is positive definite. Therefore, ρ_1 is an absolutely irreducible representation into the Lie algebra of the unitary group $U(V_1/\boldsymbol{K}, \hat{F}_1)$ (which is compact) of the quaternionic hermitian space $(V_1/\boldsymbol{K}, \hat{F}_1)$ and ρ_2 is an absolutely irreducible representation into $(II)_{p_2}$ defined by the quaternionic anti-hermitian space $(V_2/\boldsymbol{K}, \hat{F}_2)$ (see 3.3).

Conversely, it is clear that the representation ρ constructed in this manner is contained in $(III)_p$. As in 2.5, we conclude from these the following

THEOREM 2 bis. *The representation ρ in Theorem 2 with $p = q = n_0 p_0$ is equivalent to the one contained in $(III)_p$, if and only if*

1^0. *ρ_0 is an absolutely irreducible representation of $\mathfrak{g}_0$ into the Lie algebra of the orthogonal group of a definite real quadratic form of n_0 variables and ρ_{i_0} is an absolutely irreducible representation of $\mathfrak{g}_{i_0}$ into $(III)_{p_0}$ satisfying (H_2), or*

2^0. *n_0 is even, ρ_0 is an absolutely irreducible representation of $\mathfrak{g}_0$ into the Lie algebra of the unitary group of a definite quaternionic hermitian form of $n_0/2$ variables and ρ_{i_0} is an absolutely irreducible representation of $\mathfrak{g}_{i_0}$ into $(II)_{p_0}$ satisfying (H_2).*

We note that, in the case 2^0, an actual representation contained in $(III)_p$ is obtained in the form

$$\mathrm{diag.}(1_{n_0p_0}, J \otimes 1_{p_0})^{-1} \rho(X)\, \mathrm{diag.}(1_{n_0p_0}, J \otimes 1_{p_0}) \qquad \text{for } X \in \mathfrak{g},$$

where ρ is a representation given by (27) and J is an $n_0 \times n_0$ matrix expressing the semi-linear transformation $\sigma_1 : x \to xj$ of V_1 in an orthonormal

basis (e. g., if one takes the basis as given in 3. 3, then $J = \begin{pmatrix} 0 & -1_{n_0/2} \\ 1_{n_0/2} & 0 \end{pmatrix}$).

By Th. 1, 2 and 2 bis our problem is completely reduced to the following one, (if we leave aside the problem of determining the representations of compact groups mentioned in Th. 2 bis):

For a given simple Lie algebra $\mathfrak{g}$ *corresponding to an irreducible symmetric domain, determine all (equivalence-classes of) absolutely irreducible representations* ρ *of it into* $(I)_{p,q}$ *satisfying* (H_2); *furthermore, in case* $p=q$, *examine whether or not* ρ *is equivalent to a representation contained in* $(III)_p$ *or* $(II)_p$.

§ 3. Determination of absolutely irreducible representations satisfying (H_2).

3. 1. We shal now consider the problem mentioned at the end of the preceding section. Let $\mathfrak{g}$ be a (non-compact) simple Lie algebra corresponding to an irreducible symmetric domain and let ρ be an absolutely irreducible representation of $\mathfrak{g}$ into $\mathfrak{g}' = (I)_{p',q'}$ satisfying (H_2). We denote by $\Delta = \{\alpha_1, \cdots, \alpha_l\}$ a fundamental system of roots of $\mathfrak{g}_C$ relative to a Cartan subalgebra $\mathfrak{h}_C \subset \mathfrak{k}_C$ and by λ_ρ the highest weight of ρ. Then, as is well-known, λ_ρ satisfies the following condition:

$$(32) \qquad \frac{2\langle\alpha_i, \lambda_\rho\rangle}{\langle\alpha_i, \alpha_i\rangle} \ (1 \leqq i \leqq l) \text{ are non-negative integers,}$$

$\langle\ \rangle$ denoting the inner product in the dual space of $\mathfrak{h}_C$, defined by the Killing form. On the other hand, in view of (7), the condition (H_2) implies

$$(33) \qquad (-1)^{-\frac{1}{2}}\lambda(H_0) = \frac{q'}{p'+q'} \text{ or } -\frac{p'}{p'+q'}$$

for all weights λ of ρ.

For convenience of the reader, we quote here the list of all possible isomorphisms between simple Lie algebras corresponding to irreducible symmetric domains. (See [4]. These isomorphisms, being taken to satisfy (H_2), are naturally contained in the list given in 3. 10.)

$$(I)_{p,q} \cong (I)_{q,p}, \quad (I)_{1,1} \cong (III)_1 \cong (IV)_1, \quad (I)_{3,1} \cong (II)_3,$$
$$(III)_2 \cong (IV)_3, \quad (I)_{2,2} \cong (IV)_4, \quad (II)_4 \cong (IV)_6.$$

Hence to avoid overlappings, one may restrict the parameters as follows:

$$(I)_{p,q}: p \geqq q \geqq 1,\ p+q \geqq 3, \quad (II)_p: p \geqq 5,$$
$$(III)_p: p \geqq 1, \qquad (IV)_p: p \geqq 5.$$

3.2. *The case* $\mathfrak{g} = (I)_{p,q}$. In the notation of 1.3, we have

$$\Delta = \{\xi_1 - \xi_2, \cdots, \xi_{n-1} - \xi_n\} \quad (n = p + q).$$

Hence, from (32), we have

$$\lambda_\rho = \sum m_i \xi_i$$

with $m_i \in \boldsymbol{Z}$, $m_1 \geqq \cdots \geqq m_{n-1} \geqq m_n = 0$ and from (33),

$$\begin{aligned}(-1)^{-\frac{1}{2}}\lambda_\rho(H_0) &= \frac{q}{p+q} \sum_{i=1}^{p} m_i - \frac{p}{p+q} \sum_{i=p+1}^{n} m_i \\ &= \frac{1}{p+q} \sum_{\substack{1 \leqq i \leqq p \\ p+1 \leqq j \leqq n}} (m_i - m_j) \\ &= \frac{q'}{p'+q'}.\end{aligned}$$

On the other hand, applying (33) to the lowest weight $\lambda_\rho' = (m_n, m_{n-1}, \cdots, m_1)$, we obtain

$$(-1)^{-\frac{1}{2}}\lambda_\rho'(H_0) = -\frac{p}{p+q} \sum_{i=1}^{q} m_i + \frac{q}{p+q} \sum_{i=q+1}^{n} m_i = -\frac{p'}{p'+q'}.$$

Subtracting this from the preceding equality and putting $r = \mathrm{Min}\{p, q\}$, we have

$$\sum_{i=1}^{r} m_i - \sum_{i=n-r+1}^{n} m_i = 1.$$

It follows that $m_1 = 1$ and hence, if $r > 1$, we have

$$\begin{aligned}\lambda_\rho &= (1, 0, \cdots \cdots, 0, 0), \text{ or} \\ &= (1, 1, \cdots \cdots, 1, 0).\end{aligned}$$

These correspond, respectively, to the identical representation of $(I)_{p,q}$ onto itself and to the representation of it onto $(I)_{q,p}$ given by

$$(I)_{p,q} \ni X \longrightarrow J_{p,q}{}^{-1}{}^t\bar{X} J_{p,q} \in (I)_{q,p},$$

both satisfying (H_2). For $p = q > 1$, they are never contained in $(III)_p$ or $(II)_p$.

Next let $r = 1$; we may suppose that $p \geqq q = 1$. Then we have

$$\lambda_\rho = (\overbrace{1, \cdots, 1}^{m}, 0, \cdots, 0) \qquad \text{with } 1 \leqq m \leqq p,$$

which corresponds to the irreducible representation $\rho = \Lambda_m$ given by the space of skew-symmetric tensors of degree m. We shall show below that these Λ_m $(1 \leqq m \leqq p)$ actually satisfy the condition (H_2).[3]

Let (V, F) be the hermitian vector-space of signature $(p, 1)$ on which $\mathfrak{g}$ operates, and let $\Lambda(V) = \sum_{m=0}^{p+1} \Lambda_m(V)$ be the exterior algebra of V, $\Lambda_m(V)$ denoting the space of skew-symmetric tensors of degree m. One extends F to a hermitian form $F^{(m)}$ on $\Lambda_m(V)$ in a natural manner, i.e. by putting

$$(34) \quad F^{(m)}(x_1 \wedge \cdots \wedge x_m,\ y_1 \wedge \cdots \wedge y_m) = \det(F(x_i, y_j)) \quad \text{for } x_i, y_i \in V,$$

which is clearly invariant under $\rho(\mathfrak{g})$. If $(e_1, \cdots, e_{p+1})$ is an orthonormal basis of V as described in 1.3, $e_{i_1 \cdots i_m} = e_{i_1} \wedge \cdots \wedge e_{i_m}$ $(i_1 < \cdots < i_m)$ form an orthonormal basis of $\Lambda_m(V)$ with

$$F^{(m)}(e_{i_1 \cdots i_m},\ e_{i_1 \cdots i_m}) = \begin{cases} 1 & \text{if } i_m \leqq p, \\ -1 & \text{if } i_m = p+1. \end{cases}$$

Thus $F^{(m)}$ $(1 \leqq m \leqq p)$ has the signature

$$p' = \binom{p}{m}, \qquad q' = \binom{p}{m-1},$$

and we have $p' : q' = p - m + 1 : m$. On the other hand, $e_{i_1 \cdots i_m}$ is an eigen-vector corresponding to the weight

$$\lambda = \xi_{i_1} + \cdots + \xi_{i_m},$$

for which we have

$$(-1)^{-\frac{1}{2}} \lambda(H_0) = \begin{cases} \dfrac{m}{p+1} = \dfrac{q'}{p'+q'} & \text{if } i_m \leqq p, \\[2ex] \dfrac{m-1-p}{p+1} = -\dfrac{p'}{p'+q'} & \text{if } i_m = p+1. \end{cases}$$

These prove that $\rho(H_0) = H_0'$. The corresponding isometry of the symmetric domain is given as follows:

$$\begin{aligned} \mathfrak{D}(V, F) \ni z = (V_+, V_-) \longrightarrow z' = &\ (\Lambda_m(V_+) \otimes 1, \Lambda_{m-1}(V_+) \otimes V_-) \\ &\in \mathfrak{D}(\Lambda_m(V), F^{(m)}), \end{aligned}$$

where, by means of the decomposition $V = V_+ + V_-$, $\Lambda(V)$ is identified in a natural manner with $\Lambda(V_+) \otimes \Lambda(V_-)$ (as a vector-space).

Finally, in the case $p' = q'$, let us examine the condition for $\rho(\mathfrak{g})$ to be

[3] To avoid a possible confusion, the reader should remember our convention of denoting by the same letter ρ the representation of the Lie group G and the corresponding one of the Lie algebra $\mathfrak{g}$, so that the meaning of $\rho(x)$ is different according as x is considered as an element of G or as an element of $\mathfrak{g}$.

contained in $(III)_{p'}$ or $(II)_{p'}$. We first note that one has $p'=q'$, if and only if $p+1$ is even and $m=\frac{1}{2}(p+1)$. In this case, define a bilinear form B on $\Lambda_m(V)\times\Lambda_m(V)$ by the relation

$$(35) \qquad x\wedge y=B(x,y)\,e_{1\cdots p+1} \qquad \text{for } x,y\in\Lambda_m(V).$$

Then, for any $g\in GL(V)$, we have $B(\Lambda_m(g)x,\Lambda_m(g)y)=\det(g)B(x,y)$, whence follows the invariance of B under $\rho(\mathfrak{g})$. It is clear that

$$(36) \qquad B(y,x)=(-1)^{m^2}B(x,y),$$

i.e. B is symmetric or alternating according as $m\equiv 0$ or 1 (2). On the other hand, we can define a semi-linear transformation σ of $\Lambda_m(V)$ by

$$(37) \qquad F^{(m)}(x,y)=\begin{cases} B(x^\sigma,y) & \text{if } B \text{ is symmetric,}\\ \sqrt{-1}\,B(x^\sigma,y) & \text{if } B \text{ is alternating;}\end{cases}$$

then clearly σ commutes with $\rho(X)$ $(X\in\mathfrak{g})$. For an oriented subset $M=(i_1,\cdots,i_m)$ of $(1,2,\cdots,p+1)$, an easy computation shows that

$$e_M{}^\sigma=(-\sqrt{-1})\,\epsilon(M^c,M)\,\eta(M)\,e_{M^c},$$

where M^c is the complement (oriented arbitrarily) of M, $\epsilon(M^c,M)$ is the sign of the permutation $\begin{pmatrix} M^c, & M\\ 1,2,\cdots, & p+1\end{pmatrix}$, $\eta(M)$ is -1 or 1 according as $p+1\in M$ or not, and $-\sqrt{-1}$ is taken when B is alternating. From this follows that $\sigma^2=(-1)^{m^2+1}1$. Thus we conclude that

1) if $m\equiv 1$ (2), i.e. $p\equiv 1$ (4), then B is alternating and $\sigma^2=1$,

2) if $m\equiv 0$ (2), i.e. $p\equiv 3$ (4), then B is symmetric and $\sigma^2=-1$.

In the respective case, it is easy to see that $\rho(\mathfrak{g})$ is contained in $(III)_{p'}$ or $(II)_{p'}$ (cf. 3.3).

3.3. *The case* $\mathfrak{g}=(II)_p$ $(p\geqq 3)$. Let us recall briefly the definitions of the domain and the Lie algebra of type $(II)_p$. Let V be a (right) vector-space of dimension p over the real quaternion-algebra $\boldsymbol{K}$, provided with a non-degenerate anti-hermitian form $\hat{F}$. As usual, we put $\boldsymbol{K}=\boldsymbol{C}+j\boldsymbol{C}$ with $j^2=-1$, $ij=-ji$. Then it is easy to see that there exists a uniquely determined hermitian form F of signature (p,p) on V, viewed as a complex vector-space of dimension $2p$, such that we have

$$(38) \qquad \hat{F}(x,y)=i(F(x,y)-jF(xj,y))$$

and

$$F(xj,yj)=-\overline{F(x,y)} \qquad \text{for all } x,y\in V;$$

this second condition is equivalent to saying that $S(x,y)=-F(xj,y)$ is symmetric. Now, the symmetric domain $\mathfrak{D}=\mathfrak{D}(V/\boldsymbol{K},\hat{F})=(II)_p$ is, by definition, the space of all p-dimensional complex subspaces V_- of V satisfying the conditions

$$S\mid V_-=0,\qquad F\mid V_-<0. \tag{39}$$

For such a V_-, putting $V_+=V_-j$, we have an orthogonal decomposition $V=V_++V_-$ of V with respect to F, so that $\mathfrak{D}$ is canonically imbedded in $\mathfrak{D}(V,F)=(I)_{p,p}$. If one takes an orthonormal basis $(e_1,\cdots,e_p)$ of V_+ and puts $e_{p+i}=e_ij$ $(1\leqq i\leqq p)$, then $(e_1,\cdots,e_{2p})$ is an orthonormal basis of V with respect to F as described in 1.3, and, in terms of this basis, $\mathfrak{D}$ is realized as a bounded domain in $\boldsymbol{C}^{p(p-1)/2}$ formed of all skew-symmetric $p\times p$ complex matrices Z with $1_p+\bar{Z}Z>0$. The corresponding Lie algebra $\mathfrak{g}$ is given by

$$\mathfrak{g}=\{\begin{pmatrix} X_1 & X_{12}\\ -\bar{X}_{12} & \bar{X}_1\end{pmatrix}\mid X_1,X_{12}\in M_p(\boldsymbol{C}),\ {}^t\bar{X}_1=-X_1,\ {}^tX_{12}=-X_{12}\},$$

and $\mathfrak{k}$, $\mathfrak{h}$, H_0 are given exactly in the same form as in the case of $(III)_p$.

Now we can take Δ as follows:

$$\Delta=\{\xi_1-\xi_2,\cdots,\xi_{p-1}-\xi_p,\xi_{p-1}+\xi_p\}.$$

Then, from (32), we get

$$\lambda_\rho=\sum_{i=1}^p m_i\xi_i$$

with

$$m_i\in\tfrac{1}{2}\boldsymbol{Z},\ m_i\equiv m_j\ (\mathrm{mod.}\,\boldsymbol{Z}),$$
$$m_1\geqq\cdots\geqq m_{p-1}\geqq|m_p|,$$

and, from (33),

$$(-1)^{-\frac{1}{2}}\lambda_\rho(H_0)=\tfrac{1}{2}\sum_{i=1}^p m_i=\frac{q'}{p'+q'},$$

which imply, for $p\geqq 6$, that

$$m_1=1,\ m_2=\cdots=m_p=0,$$
$$p'=q'.$$

This is nothing but the identical representation of $(II)_p$ into $(I)_{p,p}$ mentioned above. $(II)_p$ is never contained in $(III)_p$.

For $3\leqq p\leqq 5$, we have, besides this identical representation, the following possibilities:

$$\begin{array}{lll} p=3, & \text{(a)} & \lambda_\rho=(\frac{1}{2},\frac{1}{2},-\frac{1}{2}),\\ & \text{(b)} & \lambda_\rho=(\frac{1}{2},\frac{1}{2},\frac{1}{2}),\\ & (*) & \lambda_\rho=(1,1,-1),\\ p=4, & \text{(c)} & \lambda_\rho=(\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}),\\ p=5, & (**) & \lambda_\rho=(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}). \end{array}$$

The representation (a) (resp. (b)) gives the isomorphism $(II)_3\cong(I)_{3,1}$ (resp. $\cong(I)_{1,3}$) (which was already given in 3.2). In the case $p=4$, we have the isomorphism $(II)_4\cong(IV)_6$, through which the identical representation and this representation (c) correspond, respectively, to the two spin representations of $(IV)_6$ satisfying (H_2) (which will be given in 3.5). The representation (*) (resp. (**)) has a weight $(-1,-1,-1)$ (rep. $(-\frac{1}{2},-\frac{1}{2},-\frac{1}{2},-\frac{1}{2},-\frac{1}{2})$) which takes value $-\frac{3}{2}i$ (resp. $-\frac{5}{4}i$) on H_0, contradicting (33). Hence the representations (*), (**) are not solutions of our problem.

3.4. *The case* $\mathfrak{g}=(III)_p$. In the notation of 1.3, we have

$$\Delta=\{\xi_1-\xi_2,\cdots,\xi_{p-1}-\xi_p,2\xi_p\}.$$

Hence, from (32), we have

$$\lambda_\rho=\sum_{i=1}^{p}m_i\xi_i$$

with $m_i\in\mathbf{Z}$, $m_1\geqq\cdots\geqq m_p\geqq 0$ and, from (33),

$$(-1)^{-\frac{1}{2}}\lambda_\rho(H_0)=\tfrac{1}{2}\sum_{i=1}^{p}m_i=\frac{q'}{p'+q'}.$$

These imply that

$$\begin{aligned} &m_1=1, m_2=\cdots=m_p=0,\\ &p'=q'. \end{aligned}$$

Thus ρ is the identical representation of $(III)_p$ into $(I)_{p,p}$. $(III)_p$ is never contained in $(II)_p$.

3.5. *The case* $\mathfrak{g}=(IV)_p$ $(p\geqq 1, p\neq 2)$. Let us recall briefly the definitions of the domain and the Lie algebra of type $(IV)_p$. Let V_R be a real vector-space provided with a non-degenerate symmetric bilinear form (or quadratic form) S of signature $(p,2)$. Call $V=V_C$ the complexification of V_R and let $x\to\bar{x}$ be the conjugation of V with respect to V_R. Then, S being extended canonically to a symmetric bilinear form on V, it is clear that

$$F(x,y)=2S(\bar{x},y)\qquad\text{for } x,y\in V \tag{40}$$

becomes a hermitian form of signature $(p, 2)$ on V. Now take an orthogonal decomposition $V = V_+ + V_-$ such that

$$\bar{V}_+ = V_+, \quad \bar{V}_- = V_-, \quad F \mid V_+ > 0, \quad F \mid V_- < 0.$$

Then, $V_{-R} = V_- \cap V_R$ is a real 2-dimensional subspace admitting (two) complex structures $\pm I$ which leaves $S \mid V_{-R}$ invariant. Therefore, putting

$$W_- = \{x \in V_- \mid Ix = ix\}, \tag{41}$$

one sees at once that W_- satisfies the following properties

$$V_- = W_- + \bar{W}_-, \tag{42}$$

$$S \mid W_- = 0, \qquad F \mid W_- < 0. \tag{43}$$

Conversely, a 1-dimensional (complex) subspace W_- of V satisfying (43) determines, by (42) and (41), a 2-dimensional (complex) subspace V_- satisfying the above conditions together with a complex structure I on V_{-R}. Now the space of all 1-dimensional subspaces W_- of V satisfying (43) is a complex manifold (submanifold of the projective space attached to V) with two connected components; and the symmetric domain $\mathfrak{D} = \mathfrak{D}(V_R, S) = (IV)_p$ is, by definition, either one of these connected components. The group G of analytic automorphisms of $\mathfrak{D}$ is given by the connected component of $PSO(V_R, S)$. If one fixes the origin $z_0 \leftrightarrow W_-^{(0)}$ and an orthonormal basis $(e_1, \cdots, e_{p+2})$ of V_R such that $f_\nu = \frac{1}{2}(e_{p+1} + \sqrt{-1}\, e_{p+2}$ generates $W_-^{(0)}$, then $\mathfrak{D} \ni z \leftrightarrow W_-$ is parametrized by $(z_1, \cdots, z_p) \in C^p$ determined by the relation

$$\sum_{i=1}^{p} e_i z_i + f_\nu + \bar{f}_\nu \zeta \in W_-,$$

where one has $\zeta = \sum_{i=1}^{p} z_i^2, |\zeta| < 1$ and $\sum_{i=1}^{p} |z_i|^2 < \frac{1}{2}(1 + |\zeta|^2)$.

In terms of the above basis $(e_1, \cdots, e_{p+2})$, the corresponding Lie algebras have the following expressions:

$$\mathfrak{g} = \{\begin{array}{c} \begin{pmatrix} \overbrace{X_1}^{p} & \overbrace{X_{12}}^{2} \\ {}^tX_{12} & X_2 \end{pmatrix} \begin{array}{l} \}p \\ \}2 \end{array} \end{array} \mid X_1, X_{12}, X_2 : \text{real}, {}^tX_i = -X_i \ (i = 1, 2)\},$$

$$\mathfrak{k} = \{\begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \in \mathfrak{g}\},$$

$$\mathfrak{p}_+ = \{\begin{pmatrix} 0 & X_{12} \\ {}^tX_{12} & 0 \end{pmatrix} \mid X_{12} \in M_{p,2}(C), X_{12}\begin{pmatrix} 1 \\ -\sqrt{-1} \end{pmatrix} = 0\},$$

$$\mathfrak{h} = \{\text{diag.}(\begin{pmatrix} 0 & -\xi_1 \\ \xi_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & -\xi_{\nu-1} \\ \xi_{\nu-1} & 0 \end{pmatrix}, (0), \begin{pmatrix} 0 & -\xi_\nu \\ \xi_\nu & 0 \end{pmatrix}) \mid \xi_i \in \boldsymbol{R}\}$$

$$(\nu = [p/2] + 1),$$

$$H_0 = \text{diag.}(0, \cdots, 0, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}).$$

Now we take a fundamenal system in the following form:

$$\Delta=\{\sqrt{-1}(\xi_\nu-\xi_1),\sqrt{-1}(\xi_1-\xi_2),\cdots,\sqrt{-1}(\xi_{\nu-2}-\xi_{\nu-1})\}$$
$$\cup\{\sqrt{-1}\,\xi_{\nu-1}\}\text{ if } p \text{ is odd, or } \{\sqrt{-1}(\xi_{\nu-2}+\xi_{\nu-1})\}\text{ if } p \text{ is even.}$$

Then the condition (32) for the highest weight becomes as follows:

$$\lambda_\rho=\sqrt{-1}\sum m_i\xi_i$$

with

$$m\in\tfrac{1}{2}\boldsymbol{Z},\quad m_i\equiv m_j\ (\mathrm{mod.}\,\boldsymbol{Z}).$$

$$m_\nu\geqq m_1\geqq\cdots\geqq m_{\nu-2}\geqq\begin{cases}m_{\nu-1}\geqq 0 & (p\colon\text{odd}),\\ |m_{\nu-1}| & (p\colon\text{even}),\end{cases}$$

and so, from (35), we have

$$(-1)^{-\frac{1}{2}}\lambda_\rho(H_0)=m_\nu-\frac{q'}{p'+q'}.$$

It follows that

$$\begin{aligned}&m_\nu=m_1=\cdots=m_{\nu-1}=\tfrac{1}{2} &&\text{for } p\colon\text{odd},\\ &m_\nu=m_1=\cdots=m_{\nu-2}=\tfrac{1}{2},\quad m_{\nu-1}=\pm\tfrac{1}{2} &&\text{for } p\colon\text{even},\\ &p'=q'.\end{aligned}$$

Thus, in either case, ρ is a spin representation of $\mathfrak{g}$.

3.6. We shall now prove that a spin representation ρ of $\mathfrak{g}=(IV)_p$ satisfies actually the condition (H_2) with respect to a certain hermitian form defined on the space of spinors.[8]

The notations being as in 3.5, denote by $C=C(V,S)$ the Clifford algebra of (V,S) (over $\boldsymbol{C}$). J (resp. ι) will represent the canonical involutorial automorphism (resp. anti-automorphism) of C determined uniquely by the property that $J\mid V=-1$ (resp. $\iota\mid V=1$). We denote by $C^\pm$ the subspaces of C formed of even and odd elements, respectively, i.e.

$$C^\pm=\{x\in C\mid x^J=\pm x\};$$

in particular, C^+ is a subalgebra. It is well-known ([2]) that, putting $\nu=[p/2]+1$, we obtain

$$C\cong\begin{cases}M_{2^\nu}(\boldsymbol{C}) & \text{for } p\colon\text{even},\\ M_{2^\nu}(\boldsymbol{C})\oplus M_{2^\nu}(\boldsymbol{C}) & \text{for } p\colon\text{odd},\end{cases}\tag{44}$$

$$C^+\cong\begin{cases}M_{2^{\nu-1}}(\boldsymbol{C})\oplus M_{2^{\nu-1}}(\boldsymbol{C}) & \text{for } p\colon\text{even},\\ M_{2^\nu}(\boldsymbol{C}) & \text{for } p\colon\text{odd}.\end{cases}\tag{45}$$

Now, consider the 'spin group'

$$G^{(1)}=\{s\in C^+_{\boldsymbol{R}}\mid s^\iota s=1,\ sV_{\boldsymbol{R}}s^{-1}=V_{\boldsymbol{R}}\},\tag{46}$$

where $C^+_{\boldsymbol{R}}=C^+\cap C_{\boldsymbol{R}}$ and $C_{\boldsymbol{R}}$ is the real Clifford algebra of $(V_{\boldsymbol{R}},S)$. For each $s\in G^{(1)}$, the inner automorphism: $x\to sxs^{-1}$ induces on $V_{\boldsymbol{R}}\subset C_{\boldsymbol{R}}$ a proper

orthogonal transformation $\phi(s)$ of (V_R, S), and the correspondence ϕ is a covering homomorphism from $G^{(1)}$ (which is connected) onto the connected component G of $SO(V_R, S)$. By (45), $G^{(1)} \subset C^+$ has two matrix representations $L^\pm$, if p is even, and one L, if p is odd, and accordingly $\rho^\pm = L^\pm \circ \phi^{-1}$ or $\rho = L \circ \phi^{-1}$ are, by definition, the spin representations of G. We have now to describe the isomorphism (45) more explicitly.

The case $p \equiv 0\ (2)$. Put

$$f_i = \tfrac{1}{2}(e_{2i-1} + \sqrt{-1}\, e_{2i}) \qquad (1 \leqq i \leqq \nu - 1),$$
$$f_\nu = \tfrac{1}{2}(e_{p+1} + \sqrt{-1}\, e_{p+2});$$

then we have

$$\begin{aligned} f_i^2 &= 0, \qquad f_i f_j = -f_j f_i, \\ \bar{f}_i f_j &= -f_j \bar{f}_i \pm^* \delta_{ij} \end{aligned} \qquad \text{for all } 1 \leqq i, j \leqq \nu, \tag{47}$$

where at the place marked with $*$ the minus sign is taken only for $i = j = \nu$. Let W_+ (resp. W_-) be the subspace of V generated by the f_i $(1 \leqq i \leqq \nu - 1)$ (resp. by f_ν) and put $W = W_+ + W_-$. Then W is a ν-dimensional subspace of V satisfying the properties

$$S \mid W = 0, \quad F \mid W \text{ has signature } (\nu - 1, 1), \tag{48}$$

and we have $V = W + \bar{W}$. Call $E = \Lambda(W)$ the exterior algebra of W, which, by (47), is identified canonically with a subalgebra of C generated by W. In the following, we use the following notations. Put $N = (1, 2, \cdots, \nu)$ and, for any oriented subset $A = (i_1, \cdots, i_a)$ of N, write $f_A = f_{i_1} \cdots f_{i_a}$. Then, from (47), one sees immediately that $E\bar{f}_N$ is a left ideal of C, which is minimal since it has dimension 2^ν. Therefore, for every element $s \in C$, there corresponds uniquely a linear transformation $L(s)$ of E such that we have

$$s x \bar{f}_N = (L(s) x) \bar{f}_N \qquad \text{for all } x \in E. \tag{49}$$

It is clear that $L(s^J) = J \circ L(s) \circ J$. It follows, in particular, that, for $s \in C^+$, $L(s)$ leaves $E^\pm = E \cap C^\pm$ invariant, inducing on each $E^\pm$ a linear transformation $L^\pm(s)$. Evidently, the couple of representations (L^+, L^-) thus obtained gives the (first) isomorphism (45).

Now we extend F to a hermitian form on E, denoted again by F, in the manner as explained in 3.2; then F satisfies the following formula:

$$F(f_A, f_B) = \begin{cases} \eta(A)\epsilon\begin{pmatrix} A \\ B \end{pmatrix} & \text{if } |A| = |B|, \\ 0 & \text{otherwise,} \end{cases} \tag{50}$$

where $|A|$ denotes the underlying set of A, $\epsilon\begin{pmatrix} A \\ B \end{pmatrix}$ is the sign of the permutation $\begin{pmatrix} A \\ B \end{pmatrix}$, and $\eta(A) = -1$ if $\nu \in |A|$ and $= 1$ otherwise. In view of

(47) and (50), one can give an alternative definition for the extension F as follows. Namely, for $x, y \in E$, one writes $\bar{x}^t y$ in the form

$$\bar{x}^t y = \sum_{A,B \subset N} f_A \bar{f}_B \gamma_{A,B}$$

with $\gamma_{A,B} \in \boldsymbol{C}$; then $F(x,y) = \gamma_{\phi,\phi}$. In other words, F is characterized by the following relation:

$$f_N{}^t \bar{x}^t y \bar{f}_N = F(x,y) f_N{}^t \bar{f}_N. \tag{51}$$

From (49) and (51), it follows at once that

$$F(L(s)x, y) = F(x, L(s^t)y), \tag{52}$$

which implies, in particular, that F is invariant under $L(s)$ $(s \in G^{(1)})$. Thus, $F' = F \mid E^+$ and $F'' = F \mid E^-$ are invariant under $L^{\pm}(s)$ $(s \in G^{(1)})$, respectively. On the other hand, one sees, from (50), that both F' and F'' have the signature $(2^{\nu-2}, 2^{\nu-2})$, according to the following orthogonal decompositions:

$$E^+ = E^+(W_+) \otimes 1 + E^-(W_+) \otimes W_-,$$
$$E^- = E^-(W_+) \otimes 1 + E^+(W_+) \otimes W_-,$$

where $E^+(W_+) = \sum_{m \equiv 0(2)} \Lambda_m(W_+)$ and $E^-(W_+) = \sum_{m \equiv 1(2)} \Lambda_m(W_+)$.

We shall now show that f_A is an eigen-vector of $\rho^{\pm}(\mathfrak{h})$ corresponding to the weight

$$\lambda = \frac{\sqrt{-1}}{2} \left(\sum_{i \notin A} \xi_i - \sum_{i \in A} \xi_i \right).$$

In fact, for $H = \mathrm{diag.}\left(\begin{pmatrix} 0 & -\xi_1 \\ \xi_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & -\xi_\nu \\ \xi_\nu & 0 \end{pmatrix} \right) \in \mathfrak{h}$, we have

$$\begin{aligned} \exp(tH) &= \mathrm{diag.}\left(\begin{pmatrix} \cos t\xi_1 & -\sin t\xi_1 \\ \sin t\xi_1 & \cos t\xi_1 \end{pmatrix}, \cdots, \begin{pmatrix} \cos t\xi_\nu & -\sin t\xi_\nu \\ \sin t\xi_\nu & \cos t\xi_\nu \end{pmatrix} \right) \\ &= \prod_{i=1}^{\nu} S_{(-\sin t\xi_i/2 e_{2i-1} + \cos t\xi_i/2 e_{2i})} \cdot S_{e_{2i}}, \end{aligned}$$

S_x $(x \in V)$ denoting the symmetry of V defined by x. Since, for $x_1, x_2 \in V$ with $S(x_1, x_1) = S(x_2, x_2) = \pm 1$, one has $x_1 x_2 \in G^{(1)}$ and $\phi(x_1 x_2) = S_{x_1} \cdot S_{x_2}$, this is

$$\begin{aligned} &= \phi\left(\prod_{i=1}^{\nu} (-\sin t\xi_i/2\, e_{2i-1} + \cos t\xi_i/2\, e_{2i}) e_{2i} \right) \\ &= \phi\left(\prod_{i=1}^{\nu-1} (\cos t\xi_i/2 - \sin t\xi_i/2\, e_{2i-1} e_{2i}) \right. \\ &\qquad \left. \times (\cos t\xi_i/2 + \sin t\xi_i/2\, e_{2i-1} e_{2i}) \right). \end{aligned}$$

Therefore, we have

$$\phi^{-1}(H)=\tfrac{1}{2}(-\sum_{i=1}^{\nu-1}\xi_i e_{2i-1}e_{2i}+\xi_\nu e_{2\nu-1}e_{2\nu})$$
$$=\sqrt{-1}\,(\sum_{i=1}^{\nu-1}\xi_i(\tfrac{1}{2}-f_i\bar{f}_i)+\xi_\nu(\tfrac{1}{2}+f_\nu\bar{f}_\nu)).$$

In view of the relation $f_i\bar{f}_i f_A\bar{f}_N=\pm f_A\bar{f}_N$ (for $i\in|A|$), $=0$ (for $i\notin|A|$), we obtain from (49)

$$L\circ\phi^{-1}(H)f_A=\frac{\sqrt{-1}}{2}\,(\sum_{i\notin A}\xi_i-\sum_{i\in A}\xi_i)f_A,$$

which proves our assertion. In particular, we have $L\circ\phi^{-1}(H_0)f_A=\pm i/2\,f_A$, according as $\nu\notin A$ or $\nu\in A$. Thus we have proved that $\rho^\pm=L^\pm\circ\phi^{-1}$ are representations of $\mathfrak{g}$ into $\mathfrak{g}'=\mathfrak{g}(E^+,F')$ and $=\mathfrak{g}(E^-,F'')$, respectively, satisfying (H_2).

Next, let us examine whether or not $\rho^\pm(g)$ are contained in $(III)_{p'}$ or $(II)_{p'}$ $(p'=2^{\nu-2})$. For that purpose, we define a bilinear form P on $E\times E$ as follows (cf. [2], 3.2). Namely, for $x,y\in E$, one writes $x^\iota y$ in the form

$$x^\iota y=\sum_{A\subset N}f_A\gamma_A$$

and put $P(x,y)=\gamma_N$; in other words, P is characterized by the following relation:

$$\bar{f}_N{}^\iota\; x^\iota y\bar{f}_N=-P(x,y)\bar{f}_N. \tag{53}$$

Then, it is clear that P is non-degenerate and we have

$$P(y,x)=(-1)^{\nu(\nu-1)/2}P(x,y), \tag{54}$$

$$P(L(s)x,y)=P(x,L(s^\iota)y)\qquad \text{for } x,y\in E, s\in C, \tag{55}$$

(55) implying, in particular, that P is invariant under $L(s)$ $(s\in G^{(1)})$. One sees also that, if ν is even, E^+ and E^- are mutually orthogonal with respect to P and $P'=P\mid E^+$ and $P''=P\mid E^-$ are invariant under $L^\pm(s)$ $(s\in G^{(1)})$, respectively.

Furthermore, define a semi-linear transformation σ of E by the formula

$$F(x,y)=\begin{cases}P(x^\sigma,y) & \text{if } P \text{ is symmetric,}\\ \sqrt{-1}\,P(x^\sigma,y) & \text{if } P \text{ is alternating.}\end{cases}$$

Then, in view of the formulas (50) and

$$P(f_A,f_B)=\begin{cases}\epsilon({}^tA,B) & \text{if } |A|^c=|B|,\\ 0 & \text{otherwise,}\end{cases}$$

where $\epsilon({}^tA,B)$ is the sign of the permutation $\begin{pmatrix}{}^tA,B\\ N\end{pmatrix}$, one gets

$$\sigma:\; f_A\longrightarrow(-\sqrt{-1})\eta(A)\epsilon({}^tA^c,A)f_{A^c},$$

where $-\sqrt{-1}$ is taken only when P is alternating. Therefore we have

$$\sigma^2 = (-1)^{\nu(\nu-1)/2+1}$$

and $(E^{\pm})^{\sigma} = E^{\pm}$, if ν is even, and $(E^{\pm})^{\sigma} = E^{\mp}$, if ν is odd. Now we distinguish two cases.

1°. $\nu \equiv 2, 3(4)$ (i.e. $p \equiv 2, 4(8)$). In this case, P is alternating and $\sigma^2 = 1$. Hence, if $\nu \equiv 2(4)$, we have $\rho^{\pm}(\mathfrak{g}) \subset (III)_{p'}$. On the contrary, if $\nu \equiv 3(4)$, we have $\rho^{-} = \sigma \circ \rho^{+} \circ \sigma \sim \bar{\rho}^{+} \sim -{}^t\rho^{+}$, so that $\rho^{\pm}$ can not be equivalent to $-{}^t(\rho^{\pm})$. Therefore, $\rho^{\pm}$ are not equivalent to any representation contained in $(III)_{p'}$ or $(II)_{p'}$.

2°. $\nu \equiv 0, 1(4)$ (i.e. $p \equiv 0, 6(8)$). In this case, P is symmetric and $\sigma^2 = -1$. Hence, if $\nu \equiv 0(4)$, we have $\rho^{\pm}(\mathfrak{g}) \subset (II)_{p'}$, while, if $\nu \equiv 1(4)$, $\rho^{\pm}$ are, as above, not equivalent to any representation contained in $(III)_{p'}$ or $(II)_{p'}$.

3.7. *The case* $p \equiv 1(2)$. In this case, let V'_R be the 2ν-dimensional (real) subspace of V_R generated by the e_i $(1 \leqq i \leqq p-1)$, e_{p+1}, e_{p+2}, and consider the correspondence $\psi \colon V' \ni x' \to x'e_p \in C^{+} = C^{+}(V, S)$. Since we have $(x'e_p)^2 = -x'^2 = -S(x', x')$, ψ can be extended uniquely to a homomorphism of $C' = C(V', -S)$ into C^{+}, which actually is an isomorphism, as is seen by comparing the dimensions. Defining f_i $(1 \leqq i \leqq \nu)$, $W, E = \Lambda(W)$ (imbedded this time in C') as before and denoting by L' the representation of C' submitted by the minimal left ideal $E \cdot \bar{f}_N$ of C', we see that $L = L' \circ \psi^{-1}$ gives the (second) isomorphism (45).

Now let F^J be the hermitian form on E extending canonically $-F$ (on W); then we have

$$F^J(x, y) = F(x^J, y) \qquad \text{for all } x, y \in E,$$

the F on the right-hand side denoting the canonical extension to E of F (on W), As before, we have

$$F^J(L'(s')x, y) = F^J(x, L'(s'^{\iota})y) \qquad \text{for } x, y \in E, s' \in C'.$$

This, together with the relations $\psi(s'^{\iota J}) = \psi(s')^{\iota}$, $L'(s'^J) = J \circ L'(s') \circ J$ $(s' \in C')$, implies the invariance of F under $L(s)$ $(s \in G^{(1)})$. Moreover, F has the signature $(2^{\nu-1}, 2^{\nu-1})$, according to the following orthogonal decomposition of E:

$$E = \Lambda(W_+) \otimes 1 + \Lambda(W_+) \otimes W_-.$$

Now, from what we have shown in the case $p \equiv 0(2)$ and from the definition of ψ, we have, for $H = \operatorname{diag.}\left(\begin{pmatrix} 0 & -\xi_1 \\ \xi_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & -\xi_\nu \\ \xi_\nu & 0 \end{pmatrix}\right)$,

$$\psi^{-1} \circ \phi^{-1}(H) = \sqrt{-1}\left(\sum_{i=1}^{\nu-1} \xi_i(\tfrac{1}{2} + f_i\bar{f}_i) + \xi_\nu(\tfrac{1}{2} - f_\nu\bar{f}_\nu)\right),$$

whence

$$L'\circ\psi^{-1}\circ\phi^{-1}(H)f_A=\frac{\sqrt{-1}}{2}\Big(\sum_{i\notin A}\xi_i-\sum_{i\in A}\xi_i\Big)f_A,$$

which proves that $\rho=L'\circ\psi^{-1}\circ\phi^{-1}$ is a representation of $\mathfrak{g}$ into $\mathfrak{g}(E,F)$ satisfying (H_2).

Finally, we examine whether or not $\rho(\mathfrak{g})$ is contained in $(III)_{p'}$ or $(II)_{p'}$ $(p'=2^{\nu-1})$. P being the bilinear form on $E\times E$ defined as before, we have

$$P(L'(s')x,y)=P(x,L'(s'^{\iota})y) \qquad \text{for } x,y\in E, s'\in C'.$$

Hence, putting

$$P^J(x,y)=P(x^J,y) \qquad \text{for } x,y\in E,$$

we see, by the same reason as above, that P^J is invariant under $L(s)$ $(s\in G^{(1)})$. From the relations

$$J^2=1, \qquad P(x^J,y^J)=(-1)^{\nu}P(x,y),$$

we have

$$P^J(y,x)=(-1)^{\nu(\nu-1)/2+\nu}P^J(x,y).$$

Furthermore, we have

$$F(x,y)=(\sqrt{-1})P^J(x^{J\circ\sigma},y)$$

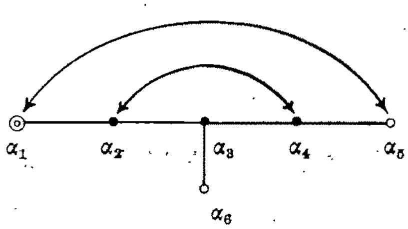
with

$$(J\circ\sigma)^2=(-1)^{\nu(\nu-1)/2+\nu+1}.$$

1°. $\nu\equiv 1,2\,(4)$ (i.e. $p\equiv 1,3\,(8)$). In this case P^J is alternating and $(J\circ\sigma)^2=1$. Hence, we have $\rho(\mathfrak{g})\subset(III)_{p'}$.

2°. $\nu\equiv 0,3\,(4)$ (i.e. $p\equiv 5,7\,(8)$). In this case, P^J is symmetric and $(J\circ\sigma)^2=-1$. Hence, we have $\rho(\mathfrak{g})\subset(II)_{p'}$.

3.8. *The case* $\mathfrak{g}=(EIII)$. Let $\Delta=\{\alpha_1,\alpha_2,\cdots,\alpha_6\}$ be a fundamental system, which we enumerate as shown in the diagram below. In particular,



α_1 will denote the (unique) 'non-compact' simple root in the sense of Harish-

Chandra[4] (cf. [4], [6]). As is well-known, this implies that we have $\alpha_1(H_0) = \pm\sqrt{-1}$ and $\alpha_i(H_0) = 0$ $(2 \leqq i \leqq 6)$; we may choose Δ in such a way that we have $\alpha_1(H_0) = \sqrt{-1}$. Then, writing λ_ρ in the form

$$\lambda_\rho = \sum_{i=1}^{6} c_i \alpha_i \qquad \text{with } c_i \in \boldsymbol{Q},$$

we have by (33)

$$(\sqrt{-1})^{-1}\lambda_\rho(H_0) = c_1 = \frac{q'}{p'+q'} \text{ or } \frac{-p'}{p'+q'}.$$

On the other hand, let $\{\varpi_1, \cdots, \varpi_6\}$ be a system of fundamental weights, enumerated in such a way that we have $\dfrac{2\langle\alpha_i, \varpi_j\rangle}{\langle\alpha_i, \alpha_i\rangle} = \delta_{ij}$. Then we can write

$$\lambda_\rho = \sum_{i=1}^{6} r_i \varpi_i$$

with non-negative integers r_i. Therefore, if we put

$$\varpi_i = \sum_{j=1}^{6} c_{ij}\alpha_j \qquad \text{with } c_{ij} \in \boldsymbol{Q},$$

we have

$$c_1 = \sum_{i=1}^{6} r_i c_{i1}.$$

Here, consulting the list of the c_{ij} ([3], 19-08), we notice that c_{i1}'s are all $\geqq 1$ except c_{51} which is $= \frac{2}{3}$. Therefore the only possibility for λ_ρ to satisfy (33) is

$$\lambda_\rho = \varpi_5 = 2/3\alpha_1 + 4/3\alpha_2 + 2\alpha_3 + 5/3\alpha_4 + 4/3\alpha_5 + \alpha_6.$$

But, denoting by w_0 the (unique) element in the Weyl group with the property $w_0\Delta = -\Delta$ and observing that $w_0 \neq -1$, we see at once that w_0 operates on Δ as follows:

$$\begin{array}{ll} \alpha_1 \leftrightarrow -\alpha_5, & \alpha_2 \leftrightarrow -\alpha_4, \\ \alpha_3 \leftrightarrow -\alpha_3, & \alpha_6 \leftrightarrow -\alpha_6. \end{array}$$

It follows that, for

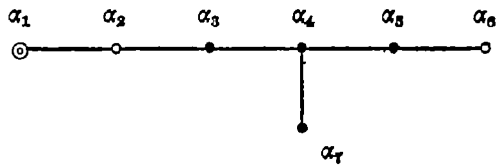
$$\lambda = w_0\varpi_5 = -4/3\alpha_1 - \cdots,$$

we have $(\sqrt{-1})^{-1}\lambda(H_0) = -4/3$, contradicting (33). Thus we conclude that there exists no solution of our problem in this case.

3.9. *The case* $\mathfrak{g} = (EVII)$. We enumerate the fundamental roots as shown in the diagram below, α_1 denoting the non-compact one.[4] Then we

---

[4] The fact that α_1 represents the non-compact root follows, for instance, from [6], Appendix, Th. (2).



notice ([3], 19-09) that, in the expressions of the fundamental weights ϖ_i $(1 \leqq i \leqq 7)$ as linear combinations of α_i $(1 \leqq i \leqq 7)$, the coefficients of α_1 are all $\geqq 1$. From this, we conclude again, by the similar arguments as in 3.8, that there exists no solution of our problem.

3.10. Summing up the results obtained in 3.2-3.9, we obtain the following list (page 461) of solutions of our problem.

For $\mathfrak{g} = (EIII)$, $(EVII)$, there is *no* solution of the problem.

UNIVERSITY OF CHICAGO.

REFERENCES.

[0] A. Borel and Harish-Chandra, "Arithmetic subgroups of algebraic groups," *Annals of Mathematics*, vol. 75 (1962), pp. 485-535.

[1] E. Cartan, "Sur les domaines bornés homogènes de l'espace de n variables complexes," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 11 (1935), pp. 116-162.

[2] C. Chevalley, *The algebraic theory of spinors*, Columbia University Press, New York, 1954.

[3] ———, "Classification des groupes de Lie algebriques," *Séminaire C. Chevalley*, ENS, 1956-58.

[4] S. Helgason, *Differential geometry and symmetric spaces*, Academic Press, New York and London, 1962.

[5] M. Kuga, *Fiber varieties over a symmetric space whose fibers are abelian varieties, I, II*, Lectures at University of Chicago, 1963-64.

[6] C. C. Moore, "Compactifications of symmetric spaces II," *American Journal of Mathematics*, vol. 86 (1964), pp. 358-378.

[7] I. Satake, "Holomorphic imbeddings of symmetric domains into a Siegel space," *Proceedings of the Conference on Complex Analysis*, University of Minnesota, 1964, Springer-Verlag, Berlin-Heidelberg-New York, 1965.

[8] C. L. Siegel, *Analytic functions of several complex variables*, Lectures at Institute for Advanced Study, 1948-49.

[9] A. Weil, "Groupes des formes quadratiques indéfinies et des formes bilinéaires alternées," *Séminaire H. Cartan*, ENS, 1957-58, Exposé No. 2.

$\mathfrak{g}$		ρ	p'	q'	Cases for which $\rho(\mathfrak{g}) \subset (\mathrm{III})_{p'}$	$\rho(\mathfrak{g}) \subset (\mathrm{II})_{p'}$
$(\mathrm{I})_{p,q}$	$p \geqq q \geqq 2$	id.	p	q		
		$\overline{\mathrm{id.}}$	q	p		
	$p \geqq q = 1$	skew-symmetric tensor representations of degree m $(1 \leqq m \leqq p)$	$\binom{p}{m}$	$\binom{p}{m-1}$	$p \equiv 1\ (4)$ $m = \frac{p+1}{2}$	$p \equiv 3\ (4)$ $m = \frac{p+1}{2}$
$(\mathrm{II})_p$	$p \geqq 5$	id.	p			always
$(\mathrm{III})_p$	$p \geqq 1$	id.	p		always	
$(\mathrm{IV})_p$	$p \geqq 4$, even	(two) spin representations	$2^{\frac{p}{2}-1}$		$p \equiv 2\ (8)$	$p \equiv 6\ (8)$
	$p \geqq 1$, odd	spin representation	$2^{\frac{p-1}{2}}$		$p \equiv 1, 3\ (8)$	$p \equiv 5, 7\ (8)$

TIGHTLY EMBEDDED 2-DIMENSIONAL POLYHEDRAL MANIFOLDS.

By Thomas F. Banchoff.*

This paper will investigate certain properties pertaining to the notion of total curvature of 2-dimensional polyhedral manifolds embedded in an n-dimensional Euclidean space E^n. The total curvature of a compact differentiable submanifold of E^n was introduced by Chern and Lashof [1] in 1957 and they defined a minimally embedded submanifold as one for which this total curvature has a minimum value.

We shall say that a 2-dimensional manifold M is *tightly embedded* in E^n if every hyperplane in E^n which contains a point of M and no nearby points is a global support hyperplane of M. For differentiable 2-manifolds, this coincides with the definition of Chern and Lashof.

In a series of papers beginning in 1959, Kuiper [2] has investigated extensively the properties of differentiable 2-manifolds satisfying this minimality or tightness condition. He has exhibited tight embeddings of the torus in E^4 and the real projective plane in E^5, and he has shown that in a certain sense this is as high as we can go in the differentiable case. Specifically he proves that if M is a differentiable compact 2-manifold tightly embedded in E^n, then M already lies in a 5-dimensional linear subspace of E^n. Since the definition of tight embedding also applies to polyhedral 2-manifolds, Kuiper asked if the polyhedral analogue of this theorem would also be true.

The major part of this paper is devoted to showing that the theory of tightly embedded polyhedral 2-manifolds is quite different from the differentiable theory, in its methods and procedures as well as in its results. We shall say that a manifold M is *substantially embedded* in E^n if it is embedded in E^n and not lying in a hyperplane. Among the results which will be proved in this paper, we mention the following theorems:

Theorem A. *For each $n \geq 3$, there exists a polyhedral 2-manifold $M(n)$ tightly and substantially embedded in E^n.*

Received July 29, 1964.

* This research was partially supported by N. S. F. grant GP-1610. The results in this paper are contained in the author's Ph. D. dissertation at the University of California, Berkeley, under Professor S. S. Chern. The author wishes to thank Professor Chern and Professor N. H. Kuiper for their encouragement and advice.

THEOREM B. *If M is a compact polyhedral 2-manifold tightly and substantially embedded in E^n for $n \geqq 6$, then*

$$n < \tfrac{1}{2}(7 + \sqrt{49 - 24\chi(M)}),$$

where $\chi(M)$ is the Euler-Poincaré characteristic of M. This is the best possible result for the torus and the real projective plane.

Numbers in brackets refer to references in the bibliography at the end of the paper.

Notations and Definitions. In the following, $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ will be a vector in E^n. The symbol $\boldsymbol{x} \cdot \boldsymbol{y}$ stands for the ordinary inner product $\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i$. Indices i, j, and k will ordinarily run from 1 to n. At $i = n$, the symbol x_{i+1} is taken to mean x_1 and at $i = 1$, x_{i-1} will mean x_n.

A compact m-dimensional manifold M embedded in E^n is the union of a finite number of cells s^i for $i = 0, 1, \cdots, m$ satisfying the following conditions:

1) s^i is a convex set included in an i-dimensional linear space in E^n and open in the relative topology. Each s^i is bounded.
2) No two distinct cells have a point in common.
3) If s is a cell in M, then the boundary of s, to be denoted bdry s, is a union of cells in M.
4) Let $St(s)$ be the union of s and all cells having s in their boundaries. Then $St(s)$ must be homeomorphic to an open convex cell in E^m.

We shall assume throughout that all of our manifolds are connected.

As is well-known, the Euler-Poincaré characteristic $\chi(M)$ of a 2-manifold M is defined to be the sum of the number of vertices and 2-cells of M diminished by the number of 1-cells.

We also define the following sets associated with a polyhedral 2-manifold embedded in E^n. The letters $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ will be vertices of M.

$$C(\boldsymbol{p}) = \{\boldsymbol{x} \in E^n \mid \boldsymbol{x} \cdot \boldsymbol{p} > \boldsymbol{x} \cdot \boldsymbol{q} \text{ for all } \boldsymbol{q} \neq \boldsymbol{p} \text{ in } M\}.$$
$$K(\boldsymbol{v}) = \{\boldsymbol{x} \in E^n \mid \boldsymbol{x} \cdot \boldsymbol{v} > \boldsymbol{x} \cdot \boldsymbol{w} \text{ for all } \boldsymbol{w} \in \operatorname{bdry} St(\boldsymbol{v})\}.$$
$$C(\boldsymbol{v}, \boldsymbol{w}) = \{\boldsymbol{x} \in E^n \mid \boldsymbol{x} \cdot \boldsymbol{v} = \boldsymbol{x} \cdot \boldsymbol{w}, \boldsymbol{x} \cdot \boldsymbol{v} > \boldsymbol{x} \cdot \boldsymbol{u} \text{ for all } \boldsymbol{u} \notin [\boldsymbol{v}\boldsymbol{w}] \text{ in } M\}.$$
$$H(\boldsymbol{v}, \boldsymbol{w}) = \{\boldsymbol{x} \in E^n \mid \boldsymbol{x} \cdot \boldsymbol{v} = \boldsymbol{x} \cdot \boldsymbol{w}\}.$$
$$H^+(\boldsymbol{v}, \boldsymbol{w}) = \{\boldsymbol{x} \in E^n \mid \boldsymbol{x} \cdot \boldsymbol{v} > \boldsymbol{x} \cdot \boldsymbol{w}\}.$$

We now make our two basic definitions:

Definition 1. A polyhedral manifold M is tightly embedded in E^n if $K(\boldsymbol{v}) \subset C(\boldsymbol{v})$ for all vertices $\boldsymbol{v}$ of M.

Definition 2. A manifold M is substantially embedded in E^n if there is no $\boldsymbol{x}$ in E^n such that $\boldsymbol{x} \cdot \boldsymbol{p}$ is constant for all $\boldsymbol{p}$ in M, other than the vector $\boldsymbol{x} = 0$.

Definition 1 states that M is tightly embedded if and only if any hyperplane containing a vertex $\boldsymbol{v}$ and no other point of $St\,\boldsymbol{v}$ must be a global support hyperplane, that is, its intersection with M is exactly the point $\boldsymbol{v}$. Definition 2 states that M is substantially embedded if it does not lie in a hyperplane orthogonal to any vector $\boldsymbol{x}$ in E^n. Thus these definitions are equivalent to those stated in the introduction.

Section 1. Construction of the Manifolds $M(n)$. We shall construct the 2-manifold $M(n)$ as a submanifold of the boundary of the n-cube, that is, the set

$$\dot{\square}^n = \{\boldsymbol{x} \in E^n \mid |x_i| \leqq 1 \text{ for all } i \text{ and } |x_j| = 1 \text{ for some } j\}$$

This set has a natural cell decomposition in which the vertices have the form $\{\boldsymbol{v} \in \dot{\square}^n \mid v_i = e_i \text{ for all } i\}$ where $e_i = 1$ or -1, 1-cells are given by $\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = e_i \text{ for all } i \neq j \text{ and } |x_j| < 1\}$, and the 2-cells have the form $\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = e_i \text{ for all } i \neq j,\ j \neq k, \text{ and } |x_j| < 1,\ |x_k| < 1\}$.

At any vertex $\boldsymbol{v}$ *of* $\dot{\square}^n$, we select the following 1-cells and 2-cells from the set $St(\boldsymbol{v})$: $\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i \text{ for } i \neq j, |x_j| < 1\}$ and $\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i \text{ for } i \neq j,\ i \neq j+1, |x_j| < 1, |x_{j+1}| < 1\}$. Now $N(\boldsymbol{v})$ denote the union of v and these 1-cells and 2-cells, $j = 1, \cdots, n$.

LEMMA 1.1. *For each $\boldsymbol{v}$ in $\dot{\square}^n$, the set $N(\boldsymbol{v})$ is homeomorphic to an open disc in E^2.*

Proof. We can give the homeomorphism explicitly by choosing n unit vectors $\boldsymbol{a}_i$ in E^2 in sequence around the unit circle not all in the same half-plane and mapping $\boldsymbol{v}$ to the origin, $\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i \text{ for } i \neq j,\ |x_j| < 1\}$ to $\{|v_j - x_j|\, \boldsymbol{a}_j\}$, and

$$\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i \text{ for } i \neq j,\ i \neq j+1,\ |x_j| < 1,\ |x_{j+1}| < 1\}$$

to $\{|v_j - x_j|\, \boldsymbol{a}_j + |v_{j+1} - x_{j+1}|\, \boldsymbol{a}_{j+1}\}$. The image of $N(\boldsymbol{v})$ under this mapping will be an open set star-shaped from the origin and therefore homeomorphic to an open disc in E^2.

We now define $M(n)$ to be the union of all the sets $N(\boldsymbol{v})$ where $\boldsymbol{v}$ ranges over the set of vertices of $\dot{\square}^n$.

LEMMA 1.2. *$M(n)$ is a polyhedral 2-manifold.*

Proof. Distinct cells in this union are disjoint, and since all vertices and 1-cells of $\dot{\square}^n$ are contained in $M(n)$, it is true that if a cell is contained in $M(n)$, then its boundary is a union of cells in $M(n)$. The set $St(s)$ for a 2-cell is just s itself. If s is the 1-cell $\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i, \text{ for } i \neq j, |x_j| < 1\}$, then $St(s)$ is the union of s and the 2-cells

$$\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i \text{ for } i \neq j-1, i \neq j, |x_{j-1}| < 1, |x_j| < 1\}$$

and

$$\{\boldsymbol{x} \in \dot{\square}^n \mid x_i = v_i \text{ for } l \neq j, i \neq j+1, |x_j| < 1, |x_{j+1}| < 1\}.$$

For a vertex $\boldsymbol{v}$ in $M(n)$, $St(\boldsymbol{v}) = N(\boldsymbol{v})$, so for every cell s in $M(n)$, $St(s)$ is homeomorphic to an open disc in E^2. Therefore $M(n)$ is a 2-dimensional polyhedral manifold.

THEOREM A. *$M(n)$ is tightly and substantially embedded in E^n.*

Proof. Let $\boldsymbol{y}$ be an element of $K(\boldsymbol{v})$ for a vertex $\boldsymbol{v}$ of $M(n)$. If $\boldsymbol{v}'$ is in $\operatorname{bdry} St(\boldsymbol{v})$, then for some j, $v_j = -v'_j$ and $v_i = v'_i$ for $i \neq j$. Then $\boldsymbol{y} \cdot \boldsymbol{v} > \boldsymbol{y} \cdot \boldsymbol{v}'$ implies that $y_j v_j > y_j v'_j$ so $y_j v_j = |y_j|$, and this will be true for each j. If $\boldsymbol{x}$ is in $M(n)$, then since $|x_i| \leqq 1$ for each i, we have $\boldsymbol{y} \cdot \boldsymbol{v} = \sum_{i=1}^{n} |y_i| \geqq \sum_{i=1}^{n} y_i x_i = \boldsymbol{y} \cdot \boldsymbol{x}$, and equality implies $\boldsymbol{x} = \boldsymbol{v}$. Therefore $\boldsymbol{y}$ is in $C(\boldsymbol{v})$. Since this is true for each vertex $\boldsymbol{v}$ of $M(n)$, we have shown that $M(n)$ is tightly embedded in E^n.

If $\boldsymbol{y} \cdot \boldsymbol{v} = \boldsymbol{y} \cdot \boldsymbol{v}'$ as above, then $y_j v_j = y_j v'_j = -y_j v_j$ so $y_j = 0$. If $\boldsymbol{y} \cdot \boldsymbol{v}$ is constant on $M(n)$, then this must be true for all vertices in $\operatorname{bdry} St(\boldsymbol{v})$ so $\boldsymbol{y}$ must be 0. Therefore $M(n)$ is substantially embedded in E^n.

Remark. Since there are n 1-cells and n 2-cells at each of the 2^n vertices of $M(n)$, we have

$$\chi(M(n)) = 2^n - \frac{n \cdot 2^n}{2} + \frac{n \cdot 2^n}{4} = 2^{n-2}(4-n).$$

Furthermore each $M(n)$ is connected and oriented.

Section 2. Substantial embeddings of polyhedral manifolds. In this section we shall prove two theorems which are true for polyhedral manifolds of arbitrary dimension although we will need only the 2-dimensional case in the proof of the results in the next section.

Before coming to the proofs of these theorems, we establish a number of lemmas.

LEMMA 2.1. $\cup \{\overline{C(\boldsymbol{v})} \mid \boldsymbol{v} \text{ in } M\} = E^n$.

Proof. If $\boldsymbol{x}$ is not in $\cup \{\overline{C(\boldsymbol{v})} \mid \boldsymbol{v} \text{ in } M\}$, then the function $\boldsymbol{p} \to \boldsymbol{x} \cdot \boldsymbol{p}$ for $\boldsymbol{p}$ in M takes its maximum at more than one vertex. Therefore $\boldsymbol{x}$ lies in the union of the hyperplanes $H(\boldsymbol{v}, \boldsymbol{w})$, each of which has measure zero in E^n. Since there are only finitely many such hyperplanes, $E^n - \cup \{C(\boldsymbol{v}) \mid \boldsymbol{v} \text{ in } M\}$ has measure zero, from which it follows that $\cup \overline{C(\boldsymbol{v})} = E^n$.

We remark that if $\boldsymbol{p}$ in M is not a vertex, then $C(\boldsymbol{p}) = \phi$.

Definition. $\mathcal{V} = \{\boldsymbol{v} \text{ in } M \mid C(\boldsymbol{v}) \neq \phi\}$.

Definition. $\mathcal{E} = \{(\boldsymbol{v}, \boldsymbol{w}) \mid \boldsymbol{v}$ and $\boldsymbol{w}$ are in M, $C(\boldsymbol{v}, \boldsymbol{w}) \neq \phi\}$.

LEMMA 2.2. *If $\boldsymbol{u}$ is in M, then there is no $\boldsymbol{x}$ in E^n such that $\boldsymbol{x} \cdot \boldsymbol{u} > \boldsymbol{x} \cdot \boldsymbol{v}$ for all $\boldsymbol{v}$ in $\mathcal{V}$.*

Proof. If $\boldsymbol{x} \cdot \boldsymbol{u} > \boldsymbol{x} \cdot \boldsymbol{v}$ for all $\boldsymbol{v}$ in $\mathcal{V}$, then by continuity, we can find an open neighborhood U of $\boldsymbol{x}$ in E^n such that $\boldsymbol{x}' \cdot \boldsymbol{u} > \boldsymbol{x}' \cdot \boldsymbol{v}$ for all $\boldsymbol{v}$ in $\mathcal{V}$ and for all $\boldsymbol{x}'$ in U. Then $U \cap \bigcup \{C(\boldsymbol{v}) \mid \boldsymbol{v} \in \mathcal{V}\} = \phi$ which contradicts Lemma 2.1, since $\bigcup\{C(\boldsymbol{v}) \mid \boldsymbol{v} \text{ in } M\} = \bigcup\{C(\boldsymbol{v}) \mid \boldsymbol{v} \text{ in } \mathcal{V}\}$.

Note that $C(\boldsymbol{v}) = \cap \{H^+(\boldsymbol{v}, \boldsymbol{w}) \mid \boldsymbol{w} \neq \boldsymbol{v}, \boldsymbol{w} \text{ in } M\}$ from the definition of $H^+(\boldsymbol{v}, \boldsymbol{w})$ as $\{\boldsymbol{x} \in E^n \mid \boldsymbol{x} \cdot \boldsymbol{v} > \boldsymbol{x} \cdot \boldsymbol{w}\}$.

LEMMA 2.3. $C(\boldsymbol{v}) = \cap \{H^+(\boldsymbol{v}, \boldsymbol{w}) \mid \boldsymbol{w} \neq \boldsymbol{v}, \boldsymbol{w} \text{ in } \mathcal{V}\}$.

Proof. If this is not true, then it follows that we can find $\boldsymbol{x}$ in $\cap \{H^+(\boldsymbol{v}, \boldsymbol{w}) \mid \boldsymbol{w} \neq \boldsymbol{v}, \boldsymbol{w} \text{ in } \mathcal{V}\}$ such that $\boldsymbol{x}$ is not in $H^+(\boldsymbol{v}, \boldsymbol{u})$ for some $\boldsymbol{u} \neq \boldsymbol{v}$ in M. Then $\boldsymbol{x} \cdot \boldsymbol{u} > \boldsymbol{x} \cdot \boldsymbol{w}$ for all $\boldsymbol{w} \neq \boldsymbol{v}$ in $\mathcal{V}$ and $\boldsymbol{x} \cdot \boldsymbol{u} > \boldsymbol{x} \cdot \boldsymbol{v}$. Let $\boldsymbol{u} - \boldsymbol{v}$ denote the difference of $\boldsymbol{u}$ and $\boldsymbol{v}$ as vectors in E^n. Then we can find $\delta > 0$ so small that $(\boldsymbol{x} + \delta(\boldsymbol{u} - \boldsymbol{v})) \cdot \boldsymbol{u} > (\boldsymbol{x} + \delta(\boldsymbol{u} - \boldsymbol{v})) \cdot \boldsymbol{w}$ for all $\boldsymbol{w} \neq \boldsymbol{v}$ in $\mathcal{V}$. But $(\boldsymbol{u} - \boldsymbol{v}) \cdot (\boldsymbol{u} - \boldsymbol{v}) > 0$ implies that we also have $(\boldsymbol{x} + \delta(\boldsymbol{u} - \boldsymbol{v})) \cdot \boldsymbol{u} > (\boldsymbol{x} + \delta(\boldsymbol{u} - \boldsymbol{v})) \cdot \boldsymbol{v}$, and this contradicts Lemma 2.2.

LEMMA 2.4. *If the open segment $(\boldsymbol{v}, \boldsymbol{w})$ is in $\mathcal{E}$, then $\boldsymbol{v}$ and $\boldsymbol{w}$ are in $\mathcal{V}$.*

Proof. Since $C(\boldsymbol{v}, \boldsymbol{w}) \neq \phi$, there is an $\boldsymbol{x}$ in E^n such that $\boldsymbol{x} \cdot \boldsymbol{v} = \boldsymbol{x} \cdot \boldsymbol{w}$ and $\boldsymbol{x} \cdot \boldsymbol{v} > \boldsymbol{x} \cdot \boldsymbol{u}$ for all $\boldsymbol{u}$ not on the closed segment $[\boldsymbol{v}, \boldsymbol{w}]$. Then as above for sufficiently small $\delta > 0$, the vector $(\boldsymbol{x} + \delta(\boldsymbol{v} - \boldsymbol{w}))$ will be in $C(\boldsymbol{v})$. Similarly $C(\boldsymbol{w}) \neq \phi$, so $\boldsymbol{v}$ and $\boldsymbol{w}$ are in $\mathcal{V}$.

THEOREM 1. *If M is substantially embedded in E^n, then we have $\nu(\mathcal{V}) \geqq n + 1$, where $\nu(\mathcal{V})$ denotes the number of elements in $\mathcal{V}$.*

Proof. If $\nu(\mathcal{V}) \leqq n$, then the points in $\mathcal{V}$ span a space of at most $n - 1$ dimensions. Thus there is an $\boldsymbol{x} \neq 0$ in E^n such that $\boldsymbol{x} \cdot \boldsymbol{v}$ is constant

for all $\boldsymbol{v}$ in $\mathcal{V}$. It follows from Lemma 2.2 that $\boldsymbol{x}\cdot\boldsymbol{v}\geqq\boldsymbol{x}\cdot\boldsymbol{u}$ for all $\boldsymbol{u}$ in M and all $\boldsymbol{v}$ in $\mathcal{V}$. Similarly $-\boldsymbol{x}\cdot\boldsymbol{v}\geqq-\boldsymbol{x}\cdot\boldsymbol{u}$, which implies that $\boldsymbol{x}\cdot\boldsymbol{v}\leqq\boldsymbol{x}\cdot\boldsymbol{u}$ for all $\boldsymbol{u}$ in M and $\boldsymbol{v}$ in $\mathcal{V}$. Thus $\boldsymbol{x}\cdot\boldsymbol{u}$ is constant for all $\boldsymbol{u}$ in M, which contradicts the hypothesis that M is substantially embedded in E^n.

Remark. This theorem follows from the fact that the convex hull of M is the same as the convex hull of $\mathcal{V}$, which is a consequence of the lemmas.

THEOREM 2. *If M is substantially embedded in E^n, then we have $n\nu(\mathcal{V})\leqq 2\nu(\mathcal{E})$.*

Proof. Let $\boldsymbol{v}$ be in $\mathcal{V}$. We shall show that at least n segments in $\mathcal{E}$ have $\boldsymbol{v}$ as an endpoint.

From Lemma 2.3, we can find a set of vertices $\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_m$ in $\mathcal{V}$ such that $C(\boldsymbol{v})=\bigcap_{i=1}^{n}\{H^+(\boldsymbol{v},\boldsymbol{w})\}$ but for each j, $D_j=\cap\{H^+(\boldsymbol{v},\boldsymbol{w}_i)\mid i\neq j\}$ strictly includes $C(\boldsymbol{v})$. We shall prove that $C(\boldsymbol{v},\boldsymbol{w}_j)\neq\phi$, and then show that $m\geqq n$.

Since $C(\boldsymbol{v})\neq\phi$, we can find $\boldsymbol{x}$ in E^n such that $\boldsymbol{x}\cdot\boldsymbol{v}>\boldsymbol{x}\cdot\boldsymbol{w}_i$ for all i. Since $C(\boldsymbol{v})\neq D_j$, we can find $\boldsymbol{y}$ in D_j such that $\boldsymbol{y}\cdot\boldsymbol{v}\leqq\boldsymbol{y}\cdot\boldsymbol{w}_j$. Then for some $\boldsymbol{z}$ on the segment $[\boldsymbol{x},\boldsymbol{y}]$, we must have $\boldsymbol{z}\cdot\boldsymbol{v}=\boldsymbol{z}\cdot\boldsymbol{w}_j$, and for any such $\boldsymbol{z}$, we will have $\boldsymbol{z}\cdot\boldsymbol{v}>\boldsymbol{z}\cdot\boldsymbol{w}_i$ for $i\neq j$. Assume that also $\boldsymbol{z}\cdot\boldsymbol{v}=\boldsymbol{z}\cdot\boldsymbol{u}$ for some $\boldsymbol{u}$ in M not equal to $\boldsymbol{v}$ or $\boldsymbol{w}_j$. If $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}_j$ are not collinear, then $H^+(\boldsymbol{u},\boldsymbol{v})\cap H^+(\boldsymbol{v},\boldsymbol{w}_j)\neq\phi$. For $\boldsymbol{z}'$ in this intersection and sufficiently small $\delta>0$, $(\boldsymbol{z}+\delta\boldsymbol{z}')\cdot\boldsymbol{v}>(\boldsymbol{z}+\delta\boldsymbol{z}')\cdot\boldsymbol{w}_i$ for $i\neq j$, and also $(\boldsymbol{z}+\delta\boldsymbol{z}')\cdot\boldsymbol{v}>(\boldsymbol{z}+\delta\boldsymbol{z}')\cdot\boldsymbol{w}_j$, so $(\boldsymbol{z}+\delta\boldsymbol{z}')$ is in $\bigcap_{i=1}^{m}\{H^+(\boldsymbol{v},\boldsymbol{w}_i)\}$, which is $C(\boldsymbol{v})$ by hypotheses. But also $(\boldsymbol{z}+\delta\boldsymbol{z}')\cdot\boldsymbol{u}>(\boldsymbol{z}+\delta\boldsymbol{z}')\cdot\boldsymbol{v}$, and this contradicts the definition of $C(\boldsymbol{v})$. Note also that if $\boldsymbol{u}$ is collinear with $\boldsymbol{v}$ and $\boldsymbol{w}_j$, then $\boldsymbol{u}$ lies between $\boldsymbol{v}$ and $\boldsymbol{w}_j$ since both $C(\boldsymbol{v})$ and $C(\boldsymbol{w}_j)$ are non-empty. Therefore $\boldsymbol{z}$ is in $C(\boldsymbol{v},\boldsymbol{w}_j)$ so $(\boldsymbol{v},\boldsymbol{w}_j)$ is in $\mathcal{E}$.

If $m<n$, then $\overline{C(\boldsymbol{v})}=\bigcap_{i=1}^{m}\{\overline{H^+(\boldsymbol{v},\boldsymbol{w}_j)}\}$ includes $\bigcap_{i=1}^{n}\{H(\boldsymbol{v},\boldsymbol{w}_j)\}$. This intersection of m hyperplanes contains a linear subspace of dimension $n-m>1$. Therefore we can find $\boldsymbol{x}\neq 0$ in E^n such that both $\boldsymbol{x}$ and $-\boldsymbol{x}$ are in $\overline{C(\boldsymbol{v})}$, so $\boldsymbol{x}\cdot\boldsymbol{v}\geqq\boldsymbol{x}\cdot\boldsymbol{u}$ and $-\boldsymbol{x}\cdot\boldsymbol{v}\geqq-\boldsymbol{x}\cdot\boldsymbol{u}$ for all $\boldsymbol{u}$ in M. But then $\boldsymbol{x}\cdot\boldsymbol{u}$ is constant for all $\boldsymbol{u}$ in M, which contradicts the hypotheses that M is substantially embedded in E^n.

Therefore for each $\boldsymbol{v}$ in $\mathcal{V}$, there are at least n segments in $\mathcal{E}$ which have $\boldsymbol{v}$ as an endpoint. By Lemma 2.4, each segment in $\mathcal{E}$ has exactly two endpoints in $\mathcal{V}$. Therefore $n\nu(\mathcal{V})\leqq 2\nu(\mathcal{E})$ and the theorem is proved.

Section 3. Tight embeddings and Euler-Poincaré characteristic; Proof of Theorem B. The major purpose of this section will be to prove Theorem B stated in the introduction. Before proceeding with this, we establish a few lemmas.

LEMMA 3.1. *If M is tightly embedded in E^n and $(\boldsymbol{v}, \boldsymbol{w})$ is in $\mathcal{E}$, then $(\boldsymbol{v}, \boldsymbol{w})$ is included in M.*

Proof. Assume that $(\boldsymbol{v}, \boldsymbol{w})$ is in $\mathcal{E}$, but that $(\boldsymbol{v}, \boldsymbol{w}) \not\subset M$. Then $C(\boldsymbol{v}, \boldsymbol{w}) \neq \phi$, so we can choose $\boldsymbol{x}$ in $C(\boldsymbol{v}, \boldsymbol{w})$. Then we can find $\boldsymbol{v}'$ and $\boldsymbol{v}''$ in $[\boldsymbol{v}, \boldsymbol{w}]$ such that $[\boldsymbol{v}, \boldsymbol{v}'] \subset M$ and $(\boldsymbol{v}', \boldsymbol{v}'') \cap M = \phi$. If $\boldsymbol{v}' = \boldsymbol{v}$, then $\boldsymbol{x}$ is in $K(\boldsymbol{v})$ but $\boldsymbol{x} \cdot \boldsymbol{v} = \boldsymbol{x} \cdot \boldsymbol{w}$ so $\boldsymbol{x}$ is not in $C(\boldsymbol{v})$, which contradicts the hypothesis that M is tightly embedded in E^n. If $\boldsymbol{v}' \neq \boldsymbol{v}$, then some $\boldsymbol{u}'$ in bdry $St(\boldsymbol{v}')$ lies in $[\boldsymbol{v}, \boldsymbol{v}']$ so $\boldsymbol{x} \cdot \boldsymbol{v} = \boldsymbol{x} \cdot \boldsymbol{u}' = \boldsymbol{x} \cdot \boldsymbol{v}' = \boldsymbol{x} \cdot \boldsymbol{w}$ and $\boldsymbol{x} \cdot \boldsymbol{v}' > \boldsymbol{x} \cdot \boldsymbol{u}$ for all $\boldsymbol{u} \neq \boldsymbol{u}'$ in bdry $St(\boldsymbol{v}')$. Then for sufficiently small $\delta > 0$, $(\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v})) \cdot \boldsymbol{v}' > (\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v})) \cdot \boldsymbol{u}$ for all $\boldsymbol{u} \neq \boldsymbol{u}'$ in bdry $St(\boldsymbol{v}')$, but we also have $(\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v})) \cdot \boldsymbol{v}' > (\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v})) \cdot \boldsymbol{u}'$ so $(\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v}))$ is in $K(\boldsymbol{v}')$. But $(\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v})) \cdot \boldsymbol{w} > (\boldsymbol{x} + \delta(\boldsymbol{w} - \boldsymbol{v})) \cdot \boldsymbol{v}'$ so $(x + \delta(\boldsymbol{x} - \boldsymbol{w}))$ is not in $C(\boldsymbol{v}')$, which again contradicts the tightness of the imbedding.

The following lemmas and the main theorem B are concerned with 2-manifolds. Therefore we asume from now on that M is a compact connected polyhedral 2-manifold and substantially embedded in E^n. Assume also that M is triangulated, so that all the 2-cells have just three sides. By Lemma 3.1 and Lemma 2.4, the union of the vertices in $\mathcal{V}$ and vertices and edges in the segments of $\mathcal{E}$ forms a closed 1-dimensional complex in M which we shall denote by $\mathcal{V} \cup |\mathcal{E}|$. Let $K_1, K_2, \cdots, K_k$ be the (non-empty) connected components of $M - (\mathcal{V} \cup |\mathcal{E}|)$. Since M is connected, no component K_i is closed. Moreover, if s is a simplex in K_i, then $St(s)$ will also be in K_i. For an arbitrary subcomplex K of M, we shall define $\chi_M(K)$ to be the number of vertices and faces of M in K diminished by the number of edges in K.

LEMMA 3.2. *If K is a connected 2-dimensional subcomplex of M which is not closed and which has the property that $St(s)$ is in K whenever s is in K, then $\chi_M(K) \leqq 1$.*

Proof. We proced by induction on the number of vertices and the number of edges in K. If K has no vertices of edges, then K consists of a single triangular face so $\chi_M(K) = 1$. Assume that the lemma is true for a subcomplex $\underline{K}$ with no vertices and less than m edges. If K has m edges and e is one of them, then e is in the boundary of exactly two triangles in K so $K - e$ falls into one or two connected components $\underline{K}_1$ and $\underline{K}_2$. ($\underline{K}_2$ may

be empty.) Each $\underline{K}_i$ fulfils the hypotheses of the lemma and has less than m edges and no vertices. Therefore

$$\chi_M(K) = \chi_M(\underline{K}_1 \cup e \cup \underline{K}_2) = \chi_M(\underline{K}_1) - 1 + \chi_M(\underline{K}_2) \leqq 1.$$

Assume now that the lemma is true for subcomplexes with less than p vertices. If K has p vertices, then since K is not closed, we can find an edge e of K such that one endpoint $\boldsymbol{v}$ is a vertex in K and the other is not. Then $K - e - \boldsymbol{v}$ satisfies the hypotheses of the lemma and has $p - 1$ vertices. Therefore $\chi_M(K) = \chi_M(K - e - \boldsymbol{v}) + \chi_M(e) + \chi_M(\boldsymbol{v}) \leqq 1$, so the lemma is proved.

LEMMA 3.3. *If M is tightly and substantially embedded in E^n, then $2\nu(\mathcal{E}) \geqq 3k$, where k is the number of components K_i.*

Proof. We first show that if $(\boldsymbol{v}, \boldsymbol{w})$ is in $\mathcal{E}$ and $(\boldsymbol{v}, \boldsymbol{w}) \cap \bar{K}_i \neq \phi$, then $(\boldsymbol{v}, \boldsymbol{w})$ is contained in $\bar{K}_i$. Assume that $\boldsymbol{u}$ in $(\boldsymbol{v}, \boldsymbol{w})$ is in $\bar{K}_i$. Since $\boldsymbol{u}$ is in $(\boldsymbol{v}, \boldsymbol{w})$, $\boldsymbol{u}$ cannot be in $\mathcal{V}$. Moreover $\boldsymbol{u}$ cannot be in any other $(\boldsymbol{v}', \boldsymbol{w}')$ in $\mathcal{E}$ since for any $\boldsymbol{x}$ in $C(\boldsymbol{v}, \boldsymbol{w})$, we have $\boldsymbol{x} \cdot \boldsymbol{u} > \boldsymbol{x} \cdot \boldsymbol{v}'$ and $\boldsymbol{x} \cdot \boldsymbol{u} > \boldsymbol{x} \cdot \boldsymbol{w}'$. Therefore exactly two edges e and e' in $St(\boldsymbol{u})$ lie in the set $\mathcal{V} \cup |\mathcal{E}|$, and these lie along the segment $[\boldsymbol{v}, \boldsymbol{w}]$. Then $St(\boldsymbol{u}) - e - e'$ falls into two pieces, one of which must be completely in $\bar{K}_i$ since K_i is connected. Therefore if $(\boldsymbol{v}, \boldsymbol{w})$ meets $\bar{K}_i$, it is contained in $\bar{K}_i$. Next, since M is connected each K_i must have at least one $\boldsymbol{v}$ in $\mathcal{V}$ in its boundary. By Theorem 2, there are at least n edges in $St(\boldsymbol{v})$ which lie along segments in $\mathcal{E}$. If we remove $\boldsymbol{v}$ and these edges from $St(\boldsymbol{v})$, then $St(\boldsymbol{v})$ falls into at least n 2-dimensional components. At least one of these components must lie in K_i, and each component is bounded by exactly two of the edges which lie along segments of $\mathcal{E}$. Therefore at least two edges in $St(\boldsymbol{v})$ lie along segments $(\boldsymbol{v}, \boldsymbol{w}_1)$ and $(\boldsymbol{v}, \boldsymbol{w}_2)$ in $\mathcal{E}$, which meet $\bar{K}_i$. Therefore by the first part of the lemma, $(\boldsymbol{v}, \boldsymbol{w}_1)$ and $(\boldsymbol{v}, \boldsymbol{w}_2)$ are contained in $\bar{K}_i$. Similarly at the vertex $\boldsymbol{w}_1$ there must be at least two segments in $\mathcal{E}$ lying in $\bar{K}_i$, so there must be altogether at least three segments of $\mathcal{E}$ in $\bar{K}_i$. Since each segment $\mathcal{E}$ lies in the boundary of at most two of the components K_i, we obtain the inequality $2\nu(\mathcal{E}) \geqq 3k$, and the lemma is proved.

We can now give the proof of Theorem B stated in the introduction:

$$\begin{aligned}\chi(M) = \chi_M(M) &= \chi_M(\mathcal{V} \cup |\mathcal{E}| \cup K_1 \cup \cdots \cup K_k) \\ &= \chi_M(\mathcal{V}) + \chi_M(|\mathcal{E}|) + \chi_M(K_1) + \cdots + \chi_M(K_k).\end{aligned}$$

But $\chi_M(\mathcal{V}) = \nu(\mathcal{V})$ and since each segment of $\mathcal{E}$ has one more edge than

vertex, $\chi_M(|\mathcal{E}|) = -\nu(\mathcal{E})$. Since $\chi_M(K_i) \leqq 1$ for each component K_i by Lemma 3.2, we obtain

$$\chi(M) \leqq \nu(\mathcal{V}) - \nu(\mathcal{E}) + k.$$

From Theorem 1 and Lemma 3.3, we then get

$$\chi(M) \leqq \frac{2}{n}\,\nu(\mathcal{E}) - \nu(\mathcal{E}) + \frac{2}{3}\,\nu(\mathcal{E}) = \frac{6-n}{3n}\,\nu(\mathcal{E}).$$

From Theorems 1 and 2, $\nu(\mathcal{E}) \geqq \dfrac{n(n+1)}{2}$ and since we assumed $n \geqq 6$ in the statement of Theorem B, we have $\dfrac{6-n}{3n} \leqq 0$. Therefore

$$\chi(M) \leqq \frac{6-n}{3n} \cdot \frac{n(n+1)}{2} = \frac{(6-n)\cdot(n+1)}{6}$$

so we obtain the quadratic inequality

$$6\chi(M) \leqq -n^2 + 5n + 6$$

or, solving this,

$$n \leqq \tfrac{1}{2}(5 + \sqrt{49 - 24\chi(M)}).$$

Adding 1 to the right hand side, we obtain the strict inequality

$$n < \tfrac{1}{2}(7 + \sqrt{49 - 24\chi(M)}).$$

This completes the proof of Theorem B.

Note. Since $\chi(M(n)) = 2^{n-2}(4-n) < \dfrac{(6-n)(n+1)}{6}$ for all $n > 3$ the bound in Theorem B is not attained by any of the manifolds $M(n)$ for $n > 3$ in Section 1.

Section 4. Tight embedding of the torus and the real projective plane. In this section we shall examine two examples which illustrate Theorem B and for which the bound in Theorem B is attained.

We shall construct an embedding of the torus T^2 as a 2-dimensional submanifold of the 6-simplex Δ^6 in E^6. Let E^6 be given as the hyperplane $\{x \in E^7 \mid \sum_{i=1}^{7} x_i = 1\}$. Then $\Delta^6 = \{x \in E^6 \subset E^7 \mid \sum_{i=1}^{7} x_i = 1 \text{ and } 0 \leqq x_i \leqq 1\}$. Consider the following triangulation of the normal form of T^2, (Figure 1), which has exactly seven vertices.

We define a map $f\colon T^2 \to \Delta^6$ which send the vertex v_j to the vertex of Δ^6 for which $x_j = 1$ and $x_i = 0$ for $i \neq j$. Since every point of T^2 is in the

interior of exactly one simplex in the above triangulation, we may extend this map linearly over all of T^2. In particular, if $\boldsymbol{x}$ in T^2 lies in the triangle with vertices $\boldsymbol{v}_i$, $\boldsymbol{v}_j$, $\boldsymbol{v}_k$ and the barycentric coordinatss of $\boldsymbol{x}$ are a_i, a_j, a_k, then $f(\boldsymbol{x})$ is the point in E^7 with i-th coordinate a_i, j-th coordinate a_j, k-th coordinate a_k, and all other coordinates 0.

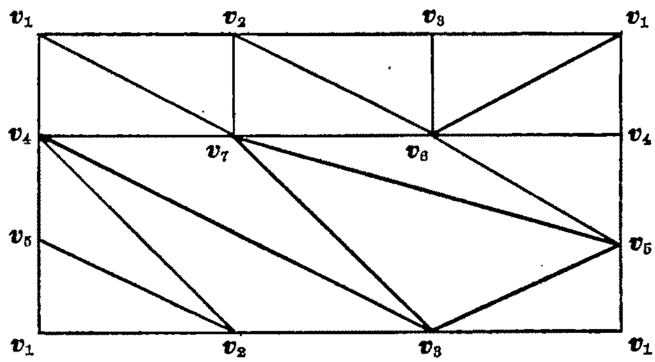


Figure 1.

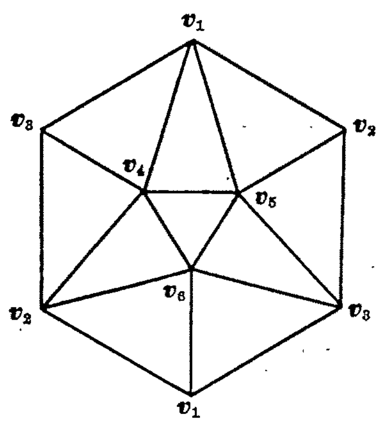


Figure 2.

By this same process, we can define a map of the real projective plane P^2 into the 5-simplex Δ^5 by making use of the following triangulation of the normal form of P^2, (Figure 2), which has six vertices.

THEOREM 3. *The embeddings of T^2 and P^2 described above are substantial and tight.*

Proof. We shall show that T^2 is substantially and tightly embedded in E^6, and the proof for P^2 in E^5 is obtained similarly.

Since $f(T^2)$ contains all the vertices of Δ^6 and Δ^6 cannot lie in a hyperplane in E^6, the fact that $f(T^2)$ is substantially embedded in E^6 follows immediately.

Since $f(T^2)$ contains all the edges of Δ^6, for any vertex $\boldsymbol{v}$ in $f(T^2)$, bdry $St(\boldsymbol{v})$ contains all the other vertices of $f(T^2)$. Therefore if $\boldsymbol{x}$ is in $K(\boldsymbol{v})$, $\boldsymbol{x}\cdot\boldsymbol{v} > \boldsymbol{x}\cdot\boldsymbol{w}$ for all $\boldsymbol{w}$ in bdry $St(\boldsymbol{v})$, so by the previous remark, $\boldsymbol{x}\cdot\boldsymbol{v} > \boldsymbol{x}\cdot\boldsymbol{w}$ for all $\boldsymbol{w} \neq \boldsymbol{v}$ in $f(T^2)$ so $\boldsymbol{x}$ is in $C(\boldsymbol{v})$. Therefore $f(T^2)$ is tightly embedded in E^6.

COROLLARY TO THEOREM 3. *The bound given in Theorem* B *for* T^2 *and for* P^2 *is the best possible.*

Proof. The Euler-Poincaré characteristic of T^2 is zero. Therefore, from Theorem B, if T^2 is tightly and substantially embedded in E^n for $n \geqq 6$, then $n < \frac{1}{2}(7 + \sqrt{49})$ so $n < 7$. But by Theorem 3, we know that T^2 can be embedded tightly and substantially in E^6, so the bound is the best possible.

Since $\chi(P^2) = 1$, Theorem B shows that if P^2 is embedded tightly and substantially in E^n for $n \geqq 6$, then $n < \frac{1}{2}(7 + \sqrt{49 - 24}) = 6$, which gives a contradiction. Therefore we cannot embed P^2 tightly and substantially in E^n for $n \geqq 6$. However, by Theorem 3, we can embed P^2 tightly and substantially in E^5, so again the bound in Theorem B is the best possible.

REFERENCES.

[1] S. S. Chern and R. K. Lashof, "On the total curvature of immersed submanifolds I," *American Journal of Mathematics*, vol. 79 (1957), pp. 306-318.

[2] N. H. Kuiper, "On convex maps," *Nieuw Archief voor Wiskunde* (3), vol. X (1962), pp. 147-164.

ON THE SURFACE OF SECTION AND PERIODIC TRAJECTORIES.

By F. Brook Fuller.

I. Introduction. This paper is concerned with the application of the Poincaré method of sections to problems of the following type. Let the autonomous differential equation $\dot{y} = F(y)$ be defined on a subset of euclidean space, with unique solutions $y(x,t)$ jointly continuous in t and the initial point x. Suppose that the domain of F contains a compact subset C with the property that the positive trajectories starting from points of C remain in C. In what circumstances is the motion $x \rightarrow y(x,t)$, $0 \leqq t < \infty$, of C into itself intercepted by a surface of section and when does C then contain periodic solutions of the equation?

A surface of section is defined in C, following Birkhoff [1], when an angular function θ can be continuously defined over C which is increasing on each trajectory, the surface of section itself being the locus $\theta = 0$. Each angular function θ on C belongs to a homotopy class γ in the set of all mappings of C into a circle. Among the homotopy classes γ Theorem 1 characterizes in the following way those which contain an angular function defining a surface of section. Take any representative angular function θ in γ and denote by $\Delta\theta(x,t)$ the net change of θ along the trajectory from x to $y(x,t)$. If $\sup \Delta\theta(x,t) > 0$ for each x in C, then for some other representative angular function θ' in γ, $\Delta\theta'(x,t)$ is an increasing function for each x, so that θ' defines a surface of section. For flows a similar theorem was proved by Schwartzman [6]. The motion of C described here is not a flow because the negative trajectories may leave C.

If each point of a surface of section S is followed out along its trajectory until its first return to S a mapping T of S into itself is defined whose periodic points lie on the periodic solutions of the differential equation. Each periodic point is a fixed point of T^n for some n. The number of fixed points of T^n, counted with multiplicities, is the Lefschetz number $\Lambda(T^n)$, determined by the endomorphisms which T induces in the homology groups of S. Theorem 2 shows that if the class γ is known to contain an angular function defining a surface of section then the numbers $\Lambda(T^n)$ can be obtained without constructing T or even knowing the differential equation, because

Received June 26, 1964.

they depend on γ alone. This is true in spite of the fact that γ does not determine the homology groups of S.

To show in a particular case that a differential equation $\dot{y} = F(y)$ has a surface of section in C by applying Theorem 1 it is necessary to select a representative angular function θ and to show $\sup \Delta\theta(x, t) > 0$. Unless a particularly fortunate choice of θ is made, the evaluation of $\Delta\theta(x, t)$ will require information about the behavior of the trajectories for large t. For this reason it is desirable to be able to demonstrate the existence of a surface of section using only information which can be calculated directly from the velocity vector F. Theorem 3 does this for the special case when C is a solid torus, where Theorems 1 and 2 are applied to show that a periodic solution exists when F satisfies the differential inequality $|\operatorname{curl} F| < k \min |F|$, the constant k depending only on the metric of C.

II. Existence of a surface section. By a *continuous one-parameter semigroup T_t acting on X* we shall mean a mapping from $X \cdot [0, \infty)$ to X which satisfies the identities $T_0(x) = x$ and $T_{s+t}(x) = T_s(T_t(x))$. By a *trajectory* of T_t we shall mean the path $T_t(x)$ obtained by fixing x and allowing t to vary from 0 to $+\infty$. By an *angular function* θ on X we shall mean a continuous function θ from X into the circle of congruence classes of real numbers reduced modulo 1. On each trajectory, if one of the representative values of θ is selected for the initial point x, θ is thereafter defined as a real number by continuation and the net change of θ along that portion of the trajectory running from x to $T_t(x)$ is a continuous real-valued function $\Delta\theta(x, t)$ on $X \cdot [0, \infty)$.

Definition. Let T_t be a continuous one-parameter semigroup acting on the compact space X. The angular function θ on X *defines a surface of section for T_t* if the net change $\Delta\theta$ of θ on each trajectory is a strictly increasing function of t. The locus $\theta \equiv 0$ is the *surface of section S* defined by θ.

THEOREM 1. *Let T_t be a continuous one-parameter semigroup acting on the compact space X. If the angular function θ on X has the property that $\Delta\theta$ is positive somewhere on each trajectory, then there is an angular function θ' in the homotopy class of θ which defines a surface of section for T_t.*

Proof. By assumption, for each x there is a t_x such that $\Delta\theta(x, t_x) > 0$. Since $\Delta\theta$ is continuous, each x has a neighborhood U_x such that $\Delta\theta(u, t_x) > p_x > 0$ for all u in U_x. Since X is compact, we may select from the $\{U_x\}$

a finite subcovering of X. Let a be the largest t_x and let p be the smallest p_x in this subcovering. Then for every x, $\Delta\theta(x,t)$ exceeds p somewhere in the interval $[0,a]$. Again by compactness there exists a b such that $\Delta\theta(x,t) > b$ throughout $[0,a]$.

The semigroup property of T_t implies that $\Delta\theta(x,t+s) = \Delta\theta(x,t) + \Delta\theta(T_t(x),s)$ so that, by induction, $\Delta\theta(x,t)$ exceeds np somewhere in the interval $[0,na]$. Now, for fixed x, if $\Delta\theta(x,m)$ is the largest value attained by $\Delta\theta$ in the interval $[0,t]$, then $0 \leqq t-m < a$ and we may write:

$$\Delta\theta(x,t) = \Delta\theta(x,m) + \Delta\theta(T_m(x),t-m) > [t/a]p + b.$$

Thus $\Delta\theta(x,t)$ converges uniformly to $+\infty$ as $t \to +\infty$.

Now define θ_λ to be the average value of θ over $[0,\lambda]$, the values to be averaged being real numbers chosen by continuation from an initial choice for $\theta(x)$. Then each θ_λ is homotopic to $\theta_0 = \theta$, λ itself being the deformation parameter. The change in θ_λ between x and $T_\epsilon(x)$ can be written in the following form:

$$\Delta\theta_\lambda = \frac{1}{\lambda}\int_0^\epsilon \Delta\theta(T_t(x),\lambda)\,dt.$$

Take λ so large that the integrand is uniformly positive. Then $\theta' = \theta_\lambda$ increases along each trajectory and so is an angular function in the homotopy class of θ which defines a surface of section, as was to be shown.

III. The periodic points of the return mapping T. When θ defines a surface of section compactness of X implies that $\Delta\theta$ increases from 0 to $+\infty$ on each trajectory, so that the equation $\Delta\theta(x,t) = +1$ has exactly one solution $t(x)$ for each x; $t(x)$ depends continuously on x. The assignment to x in S of $T_{t(x)}(x)$ in S defines the *return mapping* T of S into itself. The periodic trajectories of T_t are in one-to-one correspondence with the disjoint cycles $T(x), T^2(x), \cdots, T^n(x) = x$ of periodic points of T.

If the surface of section is a compact A. N. R., then the sum of its Betti numbers is finite and a fixed point of T^n must exist if the Lefschetz number $\Lambda(T^n) = \sum(-1)^p \operatorname{trace} T^n{}_{*p} \neq 0$, where T_{*p} is the endomorphism of the singular p-dimensional homology group with rational coefficients. We henceforth assume X to be a compact A. N. R., but an example shows that the surface of section S may not be an A. N. R. However, between a closed subset of a compact A. N. R. and any neighborhood of it one may always interpose a compact A. N. R. In our case, then, we interpose a compact A. N. R. $\mathbf{S}$ between S and the open set $\theta \not\equiv \frac{1}{2}$, whereupon the return mapping T extends, by assigning to each x in $\mathbf{S}$ the point on its trajectory for which

$\theta \equiv 0$ and $\frac{1}{2} < \Delta\theta < \frac{3}{2}$, to a mapping $\boldsymbol{T}: \boldsymbol{S} \to S \subseteq \boldsymbol{S}$. The Lefschetz number $\Lambda(\boldsymbol{T}^n)$ is easily seen to equal $\Lambda(T^n)$, while the fixed points of $\boldsymbol{T}^n$ are the same as those of T^n.

THEOREM 2. *Let T_t be a continuous one-parameter semigroup acting on the compact connected* A. N. R. *X, let the angular function θ define a surface of section for T_t and let θ_* be the induced homomorphism of the fundamental group of X into that of the circle. The Lefschetz number $\Lambda(T^n)$ of the n-th iterate of the return mapping T is determined by the homotopy class of θ in the following way: Let X' be the covering space of X corresponding to the kernel of θ_*; X' is the least covering on which θ can be continuously defined as a real number. The covering transformations of X' form a cyclic group generated by a homeomorphism g for which $\theta(g(x')) = \theta(x') + k$, where k is the index of the image of θ_* in the fundamental group of the circle. $\Lambda(T^n)$ can now be expressed as follows: $\Lambda(T^n) = k\Lambda(g^{-n/k})$ if k divides n; $\Lambda(T^n) = 0$ otherwise. The expression for $\Lambda(T^n)$ is meaningful because the assumption that θ defines a surface of section implies that k is finite and that the sum of the Betti numbers of X' is finite, even though X' is not compact.*

Proof. The covering X' will exist as described because X is a connected A. N. R. Since the fundamental group of the circle is cyclic k will be finite unless θ_* is constant, in which case θ could be continuously defined as a real number on X itself, so that $\Delta\theta(x,t)$ would be bounded, contradicting the assumption that θ defines a surface of section. The proof now breaks up into the two cases $k = 1$ and $k > 1$.

Case 1. θ_* is onto $(k = 1)$. The group of covering transformations of X' is isomorphic to the fundamental group of X modulo the kernel of θ_*, thus also to the fundamental group of the circle, and in such a way that it is generated by a g satisfying $\theta(g(x')) = \theta(x') + 1$, where we use θ again to denote the continuous real-valued function on X' covering the given θ on X.

Let $\boldsymbol{S}$ be a compact A. N. R. containing the surface of section $\theta \equiv 0$, but contained in the set $\theta \not\equiv \frac{1}{2}$, as described previously. $\boldsymbol{S}$ is covered by a homeomorphic copy $\boldsymbol{S}'$ lying in the region $-\frac{1}{2} < \theta < \frac{1}{2}$ of X'. Denote by λ the homeomorphism of $\boldsymbol{S}$ onto $\boldsymbol{S}'$ inverse to the projection of $\boldsymbol{S}'$ onto $\boldsymbol{S}$.

The semigroup T_t on X is covered by a semigroup T'_t on X', with trajectories covering those of T_t. Now define for each integer n a mapping T'_n of X' into itself as follows: $T'_n(x') = x'$ if $\theta(x') \geqq n$; if $\theta(x') \leqq n$ let

$T'_n(x')$ be the point on the trajectory from x' where $\theta = n$. Each T'_n is a deformation of X'. Furthermore, T'_n is related to $\boldsymbol{T}^n$ by the following equation: $g^{-n}T'_n\lambda(x) = \lambda\boldsymbol{T}^n(x)$ for $n \geqq 1$ and all x in $\boldsymbol{S}$.

Denote by $j: \boldsymbol{S} \to X'$ the composition of $\lambda: \boldsymbol{S} \to \boldsymbol{S}'$ with the inclusion of $\boldsymbol{S}'$ into X'. We shall show that the induced $j_*: H_p(\boldsymbol{S}) \to H_p(X')$ is onto. Denote by I_n the vector subspace $g^n{}_*j_*H_p(\boldsymbol{S})$. Any p-cycle in the region $\theta < n$ is deformed by T'_n into a cycle of $g^n\boldsymbol{S}'$ so that its homology class belongs to I_n. Thus the nested sequence of subspaces $I_0 \subseteq g_*I_0 \subseteq g^2{}_*I_0 \subseteq \cdots$ covers $H_p(X')$. But since $\boldsymbol{S}$ is an A. N. R., the subspaces are all finite-dimensional, equal to each other and to $H_p(X')$. Thus j_* is onto and the sum of the Betti numbers of X' is finite.

We now show that the kernel of j_* is the subspace M of elements m in $H_p(\boldsymbol{S})$ such that $\boldsymbol{T}^n{}_*m = 0$ for some n. The equation $g^{-n}T'_n\lambda(x) = \lambda\boldsymbol{T}^n(x)$ implies $g^{-n}T'_nj = j\boldsymbol{T}^n: \boldsymbol{S} \to X'$ and so $g^{-n}{}_*j_* = j_*\boldsymbol{T}^n{}_*: H_p(\boldsymbol{S}) \to H_p(X')$ since T' is a deformation. Thus $\boldsymbol{T}^n{}_*m = 0$ implies that $j_*m = 0$ since g_* is an automorphism. Now suppose conversely that $j_*u = 0$ is given. Let the p-cycle $\bar{u}$ be a representative of u. $j_\#\bar{u} = \partial c$ for some $(p+1)$-chain c in X'. Since c is finite, there is an n such that c belongs to the region $\theta < n$. Then $\lambda_\#\boldsymbol{T}^n{}_\#\bar{u} = g^{-n}{}_\#T'_{n\#}j_\#\bar{u} = g^{-n}{}_\#T'_{n\#}\partial c = \partial[g^{-n}{}_\#T'_{n\#}c] \sim 0$ in $\boldsymbol{S}'$. But since λ is a homeomorphism, $\boldsymbol{T}^n{}_\#\bar{u} \sim 0$ in $\boldsymbol{S}$, and we have shown that the kernel of j_* is M.

Since $\boldsymbol{T}_*(M) \subseteq M$, we may define the transformation $\boldsymbol{T}_*/M$ of $H_p(\boldsymbol{S})/M$ into itself. In a triangular matrix representation of $\boldsymbol{T}_*$ over the complex numbers the subspace M corresponds to zero diagonal entries, thus $\boldsymbol{T}_*$ and $\boldsymbol{T}_*/M$ have the same non-zero complex eigenvalues. Further, under the isomorphism of $H_p(\boldsymbol{S})/M$ with $H_p(X')$ induced by j_*, $\boldsymbol{T}_*/M$ corresponds to $g_*{}^{-1}$. Thus, for each dimension p, $\boldsymbol{T}_*$ and $g_*{}^{-1}$ have the same non-zero eigenvalues, $\boldsymbol{T}^n{}_*$ and $g^{-n}{}_*$ have the same trace and so $\Lambda(\boldsymbol{T}^n) = \Lambda(g^{-n})$, which is the conclusion of the theorem for the case $k = 1$.

Case 2. $k > 1$. In this case θ can be expressed in the form $\theta(x) = k\phi(x)$, with the consequence that ϕ_* is onto. To construct ϕ, let x_0 be any point of X, let p be a path from x_0 to x and put $\phi(x, p) = \frac{1}{k}\Delta_p\theta$, where $\Delta_p\theta$ is the net change of θ along p. For a different path q from x_0 to x,

$$\phi(x, p) = \phi(x, p) + \frac{1}{k}\Delta_{qp^{-1}}\theta,$$

but since qp^{-1} is a closed path, $\Delta_{qp^{-1}}\theta$ is a multiple of k, so that $\phi(x, p)$

defines an angular function $\phi(x)$. By putting $\phi_i = \phi + \frac{i}{k}$, $1 \leqq i \leqq k$, k solutions of $\theta = k\phi$ are obtained, each defining a surface of section for T_t.

Now the surface of section $\theta \equiv 0$ breaks into k disjoint pieces, $S_1, S_2, \cdots, S_k$, where S_i is the surface of section $\phi_i \equiv 0$, while the return mapping T permutes the pieces cyclically in such a way that $T^k(x) = T_i(x)$ on S_i, T_i being the return mapping for ϕ_i. Thus T^n has no fixed points if k does not divide n, while the fixed points of T^{rk} lying in S_i are those of T_i, so that $\Lambda(T^{rk}) = \sum_1^k (T_i^r) = k\Lambda(g^{-r})$, by applying the result of Case 1 to each ϕ_i. The description of g in the statement of the theorem is correct since $\phi_i(g(x')) = \phi_i(x') + 1$ implies $\theta(g(x')) = \theta(x') + k$ and so Theorem 2 is proved in all cases.

COROLLARY. *If the Euler characteristic of X' is not zero, then T has a periodic point and so T_t has a periodic trajectory. The period of T does not exceed the larger of $k \sum R_{2q}$ and $k \sum R_{2q+1}$, where R_p denotes the p-th Betti number of X'.*

Proof. The conclusion follows from the algebraic reasoning on the automorphism g_*^{-1} given in [3].

IV. A sufficient condition for the existence of a surface of section in a solid torus.

THEOREM 3. *Let R denote the n-dimensional solid torus with coordinates $x_1, x_2, \cdots, x_{n-1}$ in the ball of radius r $(x_1^2 + x_2^2 + \cdots + x_{n-1}^2 \leqq r^2)$ and an angular coordinate x_n $(x_n \equiv x_n + 1)$. Suppose that R has the euclidean metric $ds^2 = dx_1^2 + \cdots + dx_n^2$. Let F be a velocity field of class C^2 on R which points into R at its boundary, so that the positive trajectories of the autonomous differential equation $\dot{x} = F(x)$ stay in R. Suppose further that F satisfies the differential inequality*

$$|\operatorname{curl} F| < \frac{1}{r} \min |F|.$$

Then the semigroup T_t defined by the solutions of the differential equation has a surface of section and a periodic trajectory over which the integral of dx_n is ± 1.

Proof. $\theta = x_n$ is an angular function on R. The behavior of $\Delta\theta(x, t)$ for $t \geqq 0$ falls into one of the following three cases:

Case 1. Either $\sup \Delta\theta(x, t) = +\infty$ for all x or $\inf \Delta\theta(x, t) = -\infty$ for

all x. Then, according to Theorem 1, either θ or $-\theta$ is homotopic to an angular function defining a surface of section.

Case 2. For all x $\sup|\Delta\theta(x,t)|=+\infty$, but for some x $\sup\Delta\theta(x,t)=+\infty$, while for others $\inf\Delta\theta(x,t)=-\infty$.

Case 3. For some x, $\Delta\theta(x,t)$ is bounded.

LEMMA 1. *In cases 2 and 3 above the following is true: Given $\epsilon>0$ there is a bounding curve C in R of length L such that*

$$\int_C F\cdot dx > L(\min|F|-\epsilon).$$

Proof. In case 2 there is a point p such that $\sup\Delta\theta(p,t)=+\infty$ while p is the limit of a sequence of points q_k such that $\inf\Delta\theta(q_k,t)=-\infty$ (or else the statement is true with $+\infty$ and $-\infty$ exchanged). Now given $N>0$ there is a t such that $\Delta\theta(p,t)>N$ and, since $\Delta\theta$ is continuous, a q from the sequence q_k such that $\Delta\theta(q,t)>N$ also. But since $\inf\Delta\theta(q,t)=-\infty$ there is another $t'>t$ such that $\Delta\theta(q,t')=0$. The length L_1 of the trajectory C_1 from q to $T_{t'}(q)$ must exceed $2N$. Since q and $T_{t'}(q)$ lie in the ball $x_n=\theta(q)$ the trajectory C_1 can be completed by a directed segment C_2 of length at most $2r$ to a closed path $C=C_1+C_2$ of length L. Now

$$\begin{aligned}\int_C F\cdot dx=\int_{C_1}F\cdot dx+\int_{C_2}F\cdot dx &\geqq L_1\min|F|-2r\max|F|\\ &\geqq L(\min|F|-2rL^{-1}(\min|F|+\max|F|))\\ &> L(\min|F|-\epsilon)\ \text{for } N, \text{ hence } L,\end{aligned}$$

sufficiently large. The curve C is bounding because the integral of dx_n over C is zero.

The proof for case 3 is similar.

LEMMA 2. *Any bounding curve of length L in R bounds a surface of area at most rL.*

Proof. The required surface is swept out by the perpendiculars dropped from the curve onto the central curve $x_1=x_2=\cdots=x_{n-1}=0$.

By Stokes' theorem the integral of $F\cdot dx$ over the curve C is equal to the integral of $\operatorname{curl}F$ over any surface bounded by C. Since by Lemma 2 this surface may be taken with area at most rL, the following inequality holds:

$$\left|\int_C F\cdot dx\right|\leqq rL\max|\operatorname{curl}F|.$$

But then $rL \max |\operatorname{curl} F| > L(\min |F| - \epsilon)$ for all $\epsilon > 0$, contradicting the hypothesis $|\operatorname{curl} F| < \frac{1}{r} \min |F|$. Thus cases 2 and 3 cannot arise and T_t has a surface of section defined by an angular function homotopic to the angular function $\pm\theta$.

The homomorphism $\pm\theta_*$ described in Theorem 2 is onto. Since the covering space T' is the product of a line and a ball it has the Betti numbers of a point so that the covering transformation g^{-1} has Lefschetz number $+1$. Thus T has a fixed point and T_t has a periodic trajectory over which the integral of dx_n is ± 1, as was to be shown.

Remark. For the case $n = 2$ the hypothesis $|\operatorname{curl} F| < \frac{1}{r} \min |F|$ could be replaced by $F \neq 0$ on R. For $n > 2$, however, the author has constructed examples [4] in which, given only the hypothesis $F \neq 0$, no surface of section exists ($n = 3$) and no periodic trajectory exists ($n \geqq 4$).

CALIFORNIA INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] G. D. Birkhoff, *Dynamical Systems*, New York, 1927.
[2] F. B. Fuller "Periodic trajectories of a one-parameter semigroup," *Bulletin of the American Mathematical Society*, vol. 69 (1963), pp. 409-410.
[3] ———, "The existence of periodic points," *Annals of Mathematics*, vol. 57 (1953), pp. 229-230.
[4] ———, "Note on trajectories in a solid torus," *Annals of Mathematics*, vol. 56 (1952), pp. 438-439.
[5] S. Lefschetz, *Differential Equations: Geometric Theory*, Interscience, New York, 1957.
[6] S. Schwartzman, "Asymptotic cycles," *Annals of Mathematics*, vol. 66 (1957), pp. 270-284.

ASYMPTOTIC APPROXIMATIONS TO QUADRATIC IRRATIONALITIES, I.

By SERGE LANG.

A recent theorem of Schmidt [2] states that for almost all (real) numbers β, the number of solutions of

$$|q\beta - p| < \frac{1}{q}$$

with integers p, q and $0 < q \leqq B$ is asymptotic to $c_1 \log B$ $(B \to \infty)$ with some number $c_1 > 0$.

One could ask whether a similar estimate is not true for algebraic numbers, and also for the numbers entering in the theory of transcendental numbers (essentially those numbers in the field generated over the rationals by values of the classical functions suitably normalized, taking algebraic closure). Similarly, one can ask for extensions to other group varieties (beside the additive group). For instance, let A be an elliptic curve defined over a number field, and P a rational point. Take a suitably normalized complex analytic isomorphism of A with a complex torus (corresponding to algebraic g_2, g_3 if A is parametrized by a $\wp$-function), and let $\| \ \|$ denote the distance from the origin. One can ask for an estimate of the q such that $\| qP \| < 1/q$, and similarly for linear combinations of several points. On abelian varieties, lifting to complex n-space gives a problem concerning linear combinations of vectors.

No result of any kind except for Schmidt's theorem seems to be known regarding such asymptotic estimates. Certain machine computations for a few of the classical numbers (e, π, $\log 2, \cdots$) have a tendency to support an affirmative answer [1]. It would seem quite difficult to prove for algebraic numbers, let alone transcendental ones. Remarkably enough, it has never been noticed that the result is true for quadratic numbers. The literature mentions mostly what seems to me a freak behaviour, namely that if β is quadratic and c is sufficiently small positive number, then $|q\beta - p| < c/q$ has only a finite number of solutions, and one gets the false impression that quadratic numbers somehow misbehave. Nevertheless:

THEOREM. *Let β be a quadratic real irrational number. Let c be a*

Received August 31, 1964.

number $\geqq 1$. *For any integer* $B > 0$, *let* $\lambda(B)$ *be the number of integers* q *such that* $|q| \leqq B$ *and*

$$0 < q\beta - p < \frac{c}{|q|}$$

for some integer p. *There exist numbers* $c_1 > 0$ *and* $c_2 > 0$ *such that for all positive integers* B *we have*

$$|\lambda(B) - c_1 \log B| \leqq c_2.$$

In other words, $\lambda(B) = c_1 \log B + O(1)$.

A similar assertion holds if we replace $q\beta - p$ by $p - q\beta$ in our inequality, and hence we also get the asymptotic estimate for the number of solutions of the inequality

$$|q\beta - p| < \frac{c}{q}$$

with $0 < q \leqq B$.

The proof will be carried out by a straightforward brute force argument. We note that the number c_1 depends on c and β, even though in Schmidt's theorem, one could take c_1 to be the same for all numbers outside a set of measure 0.

We shall now prove the theorem.

Let D be an integer > 1, square free, and let

$$\alpha = \begin{cases} \sqrt{D} & \text{if } D \equiv 2, 3 \pmod 4 \\ \dfrac{1 + \sqrt{D}}{2} & \text{if } D \equiv 1 \pmod 4). \end{cases}$$

Then 1, α form a basis of the algebraic integers of $\boldsymbol{Q}(\alpha)$ over $\boldsymbol{Z}$. We let $\beta = a\alpha + b$, with rational a, b and $a \neq 0$. We let d be a positive integer such that da and db are both integers. Then $d\beta$ is an algebraic integer. We let $\bar{\beta}$ be the conjugate of β over $\boldsymbol{Q}$.

We denote by N the norm from $\boldsymbol{Q}(\alpha)$ to $\boldsymbol{Q}$.

We shall use the phrase "sufficiently large (resp. small)" to mean "greater (resp. smaller) than a constant depending only on α, β, c".

LEMMA 1. *There exists an integer* $k > 0$ *having the following property. An integer* q (*with* $|q|$ *sufficiently large*) *is such that*

$$(*) \qquad 0 < q\beta - p < \frac{c}{|q|}$$

for some p, if and only if there exists p such that $q\beta - p$ is positive, sufficiently small, and

(**) $$|N(q\beta - p)| \leqq \frac{k}{d^2}.$$

Proof. We have to distinguish cases, depending on whether $cd^2(\bar{\beta} - \beta)$ is or is not an integer. (It could be an integer only if c is equal to $t\sqrt{D}$ for some rational number t.)

Suppose that $cd^2(\bar{\beta} - \beta)$ is not an integer. Then we take

$$k = [cd^2(\bar{\beta} - \beta)].$$

First assume (*). Then

$$|N(qd\beta - dp)| < \frac{cd^2}{|q|}|q\bar{\beta} - p| = cd^2 \left|\bar{\beta} - \frac{p}{q}\right|.$$

If p/q is close to β then $\bar{\beta} - p/q$ is close to $\bar{\beta} - \beta$. Since the norm of an algebraic integer is an integer, we conclude that

$$|N(qd\beta - dp)| \leqq k,$$

thereby proving (**).

Secondly, assume (**), and also that $q\beta - p$ is positive sufficiently small. Then

$$0 < q\beta - p \leqq \frac{k}{d^2|q\bar{\beta} - p|}$$
$$\leqq \frac{c}{|q|} \frac{k}{cd^2\left|\bar{\beta} - \frac{p}{q}\right|}.$$

The quotient $k/cd^2|\bar{\beta} - \beta|$ is a fixed number < 1. For $|q|$ sufficiently large, it is clear that the right hand side of our inequality is $< c/|q|$, thereby proving (*).

Suppose that $cd^2(\bar{\beta} - \beta)$ is an integer. We must distinguish two subcases, depending on whether $\beta < \bar{\beta}$ or $\bar{\beta} < \beta$.

If $\beta < \bar{\beta}$, then for $|q|$ sufficiently large we have

$$\frac{p}{q} < \beta < \bar{\beta}.$$

We take k as before, and the first part of the argument runs as before. To conclude the argument in the second part, we now use the fact that

$$\bar{\beta} - \frac{p}{q} > \bar{\beta} - \beta.$$

If $\bar{\beta} < \beta$, then for $|q|$ sufficiently large, we have

$$\bar{\beta} < \frac{p}{q} < \beta.$$

This time, we take $k = cd^2 |\bar{\beta} - \beta| - 1$. In the first part of the argument, we have

$$|\bar{\beta} - \frac{p}{q}| < |\beta - \beta|$$

and the desired conclusion follows. The second part of the argument is carried out as before, thereby proving the lemma.

In view of the lemma, we are reduced to counting the number of integers q such that there exists p for which $q\beta - p$ is positive, sufficiently small, and

$$|N(qd\beta - dp)| \leqq k.$$

Let m be an integer, $1 \leqq m \leqq k$. We shall prove that our asymptotic estimate holds for the number of solutions of

$$|N(qd\beta - dp)| = m$$

with $q\beta - p$ positive, sufficiently small, provided that there exists at least one solution. Adding up these estimates for $m = 1, \cdots, k$ and using the fact that our original inequality actually has infinitely many solutions, we obviously obtain a proof of our theorem.

Our final step is to reduce our problem to counting certain units. Let ξ, ξ' be two algebraic integers in $\boldsymbol{Q}(\alpha)$. We say that they are *equivalent* if there exists a positive unit u such that $u\xi = \xi'$. If ξ, ξ' are equivalent then

$$|N(\xi)| = |N(\xi')|.$$

Furthermore, there is only a finite number of equivalence classes of algebraic integers in $\boldsymbol{Q}(\alpha)$ having a given norm. To prove our theorem, it will suffice to prove that the number of algebraic integers ξ satisfying the following conditions has the desired asymptotic estimate.

(1) ξ lies in a given equivalence class.

(2) ξ is positive.

(3) There exist integers q, p such that $\xi = qd\beta - dp$.

(4) ξ is sufficiently small.

LEMMA 2. *Let q_0, p_0 be integers, $q_0 \neq 0$, and let $\xi_0 = q_0 d\beta - dp_0$. The*

set of units u such that $u\xi_0$ can be written in the form $qd\beta - dp$ with integers q, p, is a group.

Proof. Let $a' = da$ and $b' = db$. We write a unit u as $u = x\alpha + y$, with integers x, y. We shall prove that the condition stated in the lemma is equivalent with a congruence condition on x.

We have

$$\xi_0 = q_0 a'\alpha + q_0 b' - dp_0.$$

Then

$$u\xi_0 = (x(q_0 b' - dp_0) + yq_0 a') \cdot \alpha + xq_0 a' D + y(q_0 b' - dp_0),$$

and we must find a necessary and sufficient condition that this expression is of type

$$qa'\alpha + qb' - dp,$$

with integers q, p. This amounts to the pair of conditions:

$$x(q_0 b' - dp_0) + yq_0 a' = qa'$$

$$xq_0 a' D + y(q_0 b' - dp_0) = qb' - dp.$$

The first one simply means that a' divides the left hand side. Let w be the g. c. d. of a' and $(q_0 b' - dp_0)$. Write $a' = wa_0$. Then the first condition is equivalent with $a_0 \mid x$ (provided $q_0 b' - dp_0 \neq 0$, a case we leave to the reader). We shall write $x = a_0 x^*$.

The first condition being satisfied, our second condition yields another divisibility condition, namely that d divides

$$x^* \left(a_0 q_0 a' D - b' \frac{q_0 b' - dp_0}{w} \right).$$

Let t be the g. c. d. of d and the expression in parentheses which we have just obtained. Write $d = td_0$. Then our last condition amounts to

$$d_0 \mid x^*.$$

Hence finally, our two conditions are equivalent with the divisibility

$$a_0 d_0 \mid x.$$

Since $u^{-1} = \pm \bar{u}$, it now follows at once that the units satisfying our divisibility condition form a group, as contended.

Let q_0, p_0 be integers, $q_0 \neq 0$, and let

$$\xi_0 = q_0 d\beta - dp_0$$

be in a given equivalence class. Assume $\xi_0 > 0$. The set of units u such that $u\xi_0$ satisfies the first three conditions (1), (2), (3) is the subgroup S of positive units in Lemma 2. Furthermore, $u\xi_0$ is sufficiently small if and only if u is sufficiently small. Hence we are reduced to counting the number of units u sufficiently small in S such that, when we write

$$u\xi_0 = qd\beta - dp$$

with integers q, p, then $|q| \leqq B$.

The group S either consists of one element, or is infinite cyclic. Assume that S is infinite. One sees at once that there exist two constants $k_1 > 0$, $k_2 > 0$ having the following property. Given any u in S, and writing $u\xi_0$ as above, we have

$$k_1 \max(|u|, |\bar{u}|) \leqq |q| \leqq k_2 \max(|u|, |\bar{u}|).$$

LEMMA 3. *There is a number $k_3 > 0$ such that the number of units u in S satisfying*

$$\max(|u|, |\bar{u}|) \leqq B$$

is equal to $k_3 \log B + O(1)$.

Proof. As usual, map a unit u in S on the vector

$$(\log |u|, \log |\bar{u}|).$$

Then S gets embedded on an infinite cyclic discrete subgroup of the straight line in the plane define by

$$\log |u| + \log |\bar{u}| = 0.$$

Our assertion is then obvious.

Let k_4 be a number > 0. We note that the number of units u in S such that

$$\max(|u|, |\bar{u}|) \leqq k_4 B$$

and the number of units u in S such that

$$\max(|u|, |\bar{u}|) \leqq B$$

differ by a bounded term. In view of the lemma and the remarks preceding it, we see that our theorem is completely proved.

We note that the number c was chosen > 1 only for definiteness. Any

$c > 0$ for which the given inequality has infinitely many solutions would do just as well. Lemma 1 and the fact that the norm of an algebraic integer must be an integer show precisely how small we can take c and still get infinitely many solutions.

Columbia University.

REFERENCES.

[1] W. Adams and S. Lang, "Some computations in diophantine approximations", *to appear.*

[2] W. Schmidt, "A metrical theorem in diophantine approximation", *Canadian Journal of Mathematics* (1959), pp. 619-631.

ASYMPTOTIC APPROXIMATIONS TO QUADRATIC IRRATIONALITIES, II.

By Serge Lang.

Let β be a (real) quadratic irrationality, and c a number $\geqq 1$. In the preceding paper, we determined asymptotically the number of integral solutions of the inequality

$$|q\beta - p| < \frac{c}{|q|}$$

for $|q| \leqq B$ and $B \to \infty$. We shall now consider the more refined inequality

$$|q\beta - p| < \psi(q)$$

where ψ is a suitable function. Again one expects that in general the number is asymptotic to

$$\Psi(B) = c_1 \int_1^B \psi(t)\, dt,$$

with some constant $c_1 > 0$, but we shall have to make a growth assumption on ψ in order to obtain this result.

When $\psi(x) = \rho$ with $0 < \rho < 1$, the problem is one of equidistribution for the numbers $q\beta$ on the circle, and has been considered before, notably by Hecke [1], who introduced the corresponding Dirichlet generating series, proved that it is meromorphic, and obtained a rather good error term. Hecke in fact uses a method which allows him to deal with any real number β such that for every $\epsilon > 0$ we have

$$|q\beta - p| > \frac{1}{q^{1+\epsilon}}$$

for all but a finite number of q, and hence in view of Roth's theorem, Hecke's theorem applies to algebraic numbers. (See the last theorem in Hecke's paper.)

Although Hecke's method works therefore rather well in his case, it cannot work for the counting problem with a function ψ, unless one knows something about the analytic behaviour of the functions which would play in the present questions a role similar to the L-functions in the theory of prime numbers.

Received October 8, 1964.

The method used in the present paper does not apply to Hecke's case as it stands, but works in other cases, for instance when

$$\psi(x) = \frac{(\log x)^m}{x}$$

with any $m > 0$. Thus our result is in a certain sense complementary to Hecke's.

Generally speaking, the study of the asymptotic distribution of field elements (as distinguished from ideals), would seem to deserve more attention than it has up to now.

1. Statement of the theorem. Let ψ be a real function of a real variable t, defined for $t > 0$, monotone decreasing to 0. We let $\omega(t) = t\psi(t)$, so $\psi(t) = \omega(t)/t$. We shall assume that ω is of class C^1, positive, strictly monotone increasing to infinity, and say that $\omega(t) = O(t^{\frac{1}{2}})$. (Our result will not apply to an ω growing too fast.)

As a matter of notation, we let f be the inverse function of ω. It is convenient to extend the domain of definition of ψ and ω to negative t by letting $\psi(t) = \psi(|t|)$, and similarly for ω.

We shall prove a theorem concerning norms:

THEOREM 1. *Let β be a real quadratic irrationality. Let β_1, β_2 be a basis of $\boldsymbol{Q}(\beta)$ over $\boldsymbol{Q}$, with β_1, β_2 algebraic integers. Let $\lambda(B)$ be the number of integral pairs (q_1, q_2) such that*

$$|q_1\beta_1 + q_2\beta_2| \leqq 1, \qquad |q_1| \leqq B$$

and

$$|N(q_1\beta_1 + q_2\beta_2)| \leqq \omega(q_1).$$

Then there exists a constant $c_1 > 0$ such that for $B \to \infty$ we have

$$\lambda(B) = c_1 \int_1^B \psi(t)\,dt + O(\omega(B) + (\log B)\omega(B)^{\frac{1}{2}}).$$

When ω does not grow too fast, the integral is usually asymptotic to a constant times $\omega(B)\log B$, so that the error term is of a lower order of magnitude. For example, if $\omega(t) = (\log t)^\rho$ with $0 < \rho < 1$, we get the asymptotic estimate

$$c_1 \frac{(\log B)^{\rho+1}}{\rho + 1} + O((\log B)^\rho + (\log B)^{(\rho/2)+1}).$$

Similarly when $\omega(t) = \log\log t$, or further iterated logs. However, if $\omega(t) = t^\delta$ with $0 < \delta \leqq 1$, our result does not assert anything because the error term is of the same order of magnitude as the main term.

We shall now show how Theorem 1 implies a theorem concerning approximations to β. Given $0 < c < 1$, there is only a finite number of $\xi = q\beta - p$ such that $c \leqq |\xi| \leqq 1$, and

$$|\bar{\xi}| \leqq \omega(q)/c.$$

Indeed, we note that $\max(|q|, |p|)$ and $\max(|\xi|, |\bar{\xi}|)$ are of the same order of magnitude. Hence as $\max(|q|, |p|) \to \infty$, it follows that $|\bar{\xi}|$ grows faster than $\omega(q)$.

Suppose that (q, p) is a pair of integers such that

$$(1) \qquad |q\beta - p| < \frac{\omega(q)}{q}.$$

Let d be an integer > 0 such that $d\beta$ is an algebraic integer. Then

$$(2) \qquad \begin{aligned} |N(qd\beta - dp)| &< \frac{\omega(q)}{|q|} d^2 |q\bar{\beta} - p| \\ &< \omega(q) d^2 \left(|\bar{\beta} - \beta| + \frac{\omega(q)}{q^2}\right). \end{aligned}$$

Let $\omega_1(t) = \omega(t) d^2 (|\bar{\beta} - \beta| + \omega(t)/t^2)$. Then ω_1 is usually also strictly increasing for sufficiently large t. A simple sufficient condition, for instance, is that $\omega'(t) > 1/t^2$, as one sees at once by taking the derivative.

The set of solutions of (1) is contained in the set of solutions of (2). If we apply Theorem 1, taking $\beta_1 = d\beta$ and $\beta_2 = d$, using the function ω_1 instead of ω, we find that the number of solutions of (2), up to the given error term, is the same as the number of solutions of

$$(3) \qquad |N(qd\beta - dp)| \leqq \omega(q) d^2 |\bar{\beta} - \beta|.$$

Conversely, let $\omega_2(t) = \omega(t) d^2 (|\bar{\beta} - \beta| - 2\omega(t)/t^2)$. Let (q, p) be a pair of integers such that

$$(4) \qquad |N(qd\beta - dp)| \leqq \omega_2(q),$$

and such that $|q\beta - p| \leqq 1$, $|q|$ is sufficiently large. Then in fact, $|q\beta - p|$ is small, and a simple computation shows that $|q\beta - p|$ satisfies (1). The function $\omega_2(t)$ is strictly increasing (no extra condition is needed this time). Thus we can apply Theorem 1 again, and find that the number of solutions of (4), up to the given error term, is the same as the number of solutions of (3). Since the desired solutions of (1) are squeezed in between, we see that Theorem 1 implies:

THEOREM 2. *Let β be a real quadratic irrationality. Let ω, ψ be as*

before, assume in addition that $\omega'(t) > 1/t^2$ for all t sufficiently large. There exists a constant $c_1 > 0$ such that the number of integral solutions (q, p) for the inequality

$$|q\beta - p| < \psi(q)$$

with $|q| \leqq B$ is equal to

$$c_1 \int_1^B \psi(t)\,dt + O(\omega(B) + (\log B)\omega(B)^{\frac{1}{2}}).$$

2. Proof of the theorem. We prove Theorem 1. Let us introduce some notation. Let L be the module generated over the integers $\boldsymbol{Z}$ by β_1, β_2. Denote by $\lambda_L(B)$ the number of elements of L satisfying the conditions stated in Theorem 2.

We embed K in $\boldsymbol{R}^2$ as usual. If $\xi \in K$, we map ξ on the vector $(\sigma\xi, \bar{\sigma}\xi)$ where σ, $\bar{\sigma}$ are the conjugate embeddings of K in $\boldsymbol{R}$. We shall identify ξ with $\sigma\xi$, and thus write this vector $(\xi, \bar{\xi})$. Under this mapping, we see that L is embedded on a lattice or rank 2 in $\boldsymbol{R}^2$.

Let $\mathfrak{o}_L$ be the subring of K consisting of all elements γ such that $\gamma L \subset L$. Then $\mathfrak{o}_L$ is a subring of the ring of algebraic integers I_K, of rank 2 over $\boldsymbol{Z}$. It is easily seen that the group of units U_L of $\mathfrak{o}_L$ is a subgroup of finite index in the group of units U_K of I_K.

By an L-ideal we shall mean a submodule $\mathfrak{a} \neq 0$ of L such that $\mathfrak{o}_L\mathfrak{a} = \mathfrak{a}$. If $\mathfrak{a}$ is a principal L-ideal, and ξ, ξ' are two generators of $\mathfrak{a}$ over $\mathfrak{o}_L$, then there exists a unit $u \in U_L$ such that $\xi' = u\xi$. We denote $\mathfrak{o}_L\xi$ by (ξ). If $\mathfrak{a} = (\xi)$, we define $\boldsymbol{N}\mathfrak{a} = |N(\xi)|$.

LEMMA 1. *The number of principal L-ideals $\mathfrak{a}$ such that $\boldsymbol{N}\mathfrak{a} \leqq B$ is equal to $c_L B + O(B^{\frac{1}{2}})$, with some constant $c_L > 0$.*

Proof. The argument is entirely similar to the classical one when $L = I_K$. A careful treatment (which simplifies considerably in the special case under consideration) for the classical case, without hand waving, will be found in Schanuel [3]. It consists in estimating the number of lattice points in a homogeneously expanding region, and we shall omit it here.

If $\xi \in L$ and $\xi = q_1\beta_1 + q_2\beta_2$, then we let $q(\xi) = |q_1|$.

LEMMA 2. *There exist constants c_2, $c_3 > 0$, depending only on L, having the following property. For any principal L-ideal $\mathfrak{a}$, the number of $\beta \in L$ such that $(\xi) = \mathfrak{a}$, $|\xi| \leqq 1$, $q(\xi) \leqq B$, and $|N(\xi)| \leqq \omega(q(\xi))$, differs from*

$$c_2(\log B - \log f(\boldsymbol{N}\mathfrak{a}))$$

by a term bounded by c_3.

Proof. Let T be the set of $\xi \in L$ such that $(\xi) = \mathfrak{a}$. We map T into $\boldsymbol{R}^2$ by the usual log mapping,

$$l: \xi \rightarrow (\log |\xi|, \log |\bar{\xi}|).$$

If $\xi_0 \in T$, then $l(T) = l(\xi_0) + l(U_L)$, and $l(U_L)$ is a discrete subgroup of the hyperplane (line) $x + y = 0$ in $\boldsymbol{R}^2$. Thus $l(T)$ is the translation by $l(\xi_0)$ of this subgroup on the line $x + y = \log N\mathfrak{a}$. If we let $l(\xi) = (x, y)$, then our conditions on ξ can be expressed by saying that

$$x \leqq 0, \qquad \log q(\xi) \leqq \log B,$$

and

$$f(N\mathfrak{a}) \leqq q(\xi) \leqq B.$$

There are two constants $k_1, k_2 > 0$ such that for all $\xi \in L$, $\xi \neq 0$ with $|\xi| \leqq 1$ we have

$$k_1 |\bar{\xi}| \leqq q(\xi) \leqq k_2 |\bar{\xi}|.$$

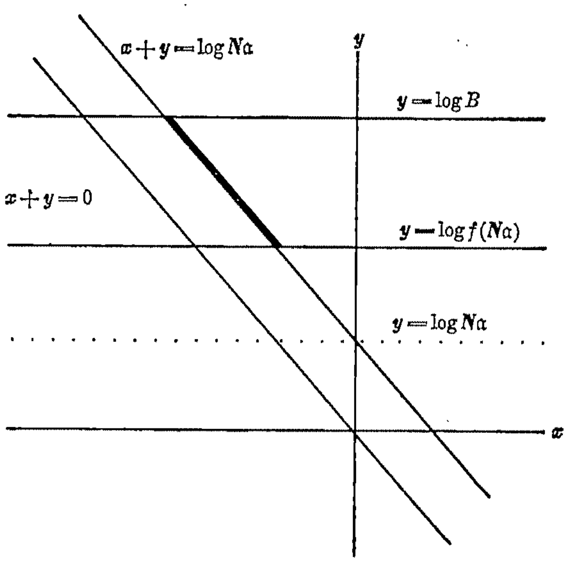
On the straight line $x + y = \log N\mathfrak{a}$, a point (x, y) of $l(T)$ is such that $y = \log |\bar{\xi}|$. The number of such points with $x < 0$ and

$$\log f(N\mathfrak{a}) + \log k_1 \leqq y \leqq \log B + \log k_2$$

is equal to

$$c_2(\log B - \log f(N\mathfrak{a})) + O(1).$$

This proves our lemma.



The above picture is a diagram of the preceding proof. One determines

the lattice points on the line $x+y=\log N\mathfrak{a}$ between $y=\log f(N\mathfrak{a})$ and $y=\log B$.

In view of Lemmas 1 and 2, our number $\lambda_L(B)$ is given by

$$c_2^{-1}\lambda_L(B)=\sum_{N\mathfrak{a}\leqq\omega(B)}(\log B-\log f(N\mathfrak{a}))+O(\omega(B)).$$

We shall conclude the proof by summing by parts to get the desired asymptotic expression.

For each positive integer ν, let $A(\nu)$ be the number of principal L-ideals $\mathfrak{a}$ such that $N\mathfrak{a}\leqq\nu$. Let us denote by $S(B)$ the sum on the right hand side of our expression for $c_2^{-1}\lambda_L(B)$. Then

$$\begin{aligned}S(B)&=\sum_{\nu=1}^{[\omega(B)]}\log\left(\frac{B}{f(\nu)}\right)(A(\nu)-A(\nu-1))\\&=(\log B-\log f([\omega(B)]))A([\omega(B)])\\&\quad+\sum_{\nu=1}^{[\omega(B)]-1}(\log f(\nu+1)-\log f(\nu))A(\nu).\end{aligned}$$

We now write $A(\nu)=c_L\nu+O(\nu^{\frac{1}{2}})$. Then, up to a constant factor c_L, the main term of our expression becomes

$$(\log B-\log f([\omega(B)]))[\omega(B)]+\sum_{\nu=1}^{[\omega(B)]-1}(\log f(\nu+1)-\log f(\nu))\nu.$$

Unwinding this sum by parts again, we obtain

$$(\log B)[\omega(B)]-\sum_{\nu=1}^{[\omega(B)]}\log f(\nu).$$

Since f and hence $\log f$ are strictly increasing, we have

$$\begin{aligned}\sum_{\nu=1}^{[\omega(B)]}\log f(\nu)&=\int_1^{\omega(B)}\log f(t)\,dt+O(\log B)\\&=(\log B)\omega(B)-\int_1^{\omega(B)}\frac{t}{f(t)}f'(t)\,dt+O(\log B),\end{aligned}$$

using integration by parts. Hence the main term in our expression is equal to

$$\int_1^{\omega(B)}\frac{t}{f(t)}f'(t)\,dt+O(\log B).$$

Changing variables, putting $t = \omega(\tau)$, we find that this main term is the desired one, i. e.

$$\int_{f(1)}^{B} \frac{\omega(t)}{t}\, dt + O(\log B).$$

There remains to estimate the error term which we had previously, namely

$$(\log B - \log f([\omega(B)]))[\omega(B)]^{\frac{1}{2}} + \sum_{\nu=1}^{[\omega(B)]-1} (\log f(\nu+1) - \log f(\nu))\nu^{\frac{1}{2}}.$$

Unwinding this sum as in the preceding case, we get

$$(\log B)[\omega(B)]^{\frac{1}{2}} + \sum_{\nu=1}^{[\omega(B)]} \log f(\nu)(\nu^{\frac{1}{2}} - (\nu-1)^{\frac{1}{2}}).$$

Each term in the sum is positive. We replace $f(\nu)$ by $f(\omega(B)) = B$, and see that our error is bounded by a constant times $(\log B)\omega(B)^{\frac{1}{2}}$, as was to be shown.

3. Remarks. One sees at once from the various constructions of the proof that $c_1 = c_2 c_L$, and hence that c_1 depends only on L, i. e. β (or β_1, β_2), and not on the auxiliary functions ψ or ω. Thus for functions ψ such that the error term is small compared to the main term, this main term depends linearly on ψ.

In the preceding paper, we consider the number of $\xi \in L$ satisfying the inequality $|N(\xi)| \leqq k$ with some constant k, equal to a positive integer. We can apply the method of proof for Theorem 1, using the units of $\mathfrak{o}_L$ (this being a slight variation of the method used in [2]). One then sees that the number of solutions of this inequality with $q(\xi) \leqq B$, $|\xi| \leqq 1$, is equal to

$$c_2 A(k) \log B + O(1),$$

where c_2 is a constant determined by the length of a fundamental domain for the units, and $A(k)$ is the number of L-ideals $\mathfrak{a}$ such that $N\mathfrak{a} \leqq k$.

From this one also sees that our restrictions on the function ω are necessary. Indeed, ω cannot tend to 0 in view of the existence of the constant $c > 0$ such that $|q\beta - p| > c/|q|$. Furthermore, ω cannot oscillate, because using Lemma 1 of the preceding paper, if we let ω oscillate very slightly near a critical value of the constant k then it follows at once that there is no asymptotic value for the number of solutions of our inequality (the

constant $A(k)$ jumps by discrete amounts). Thus in dealing with quadratic irrationalities, one must assume that ω is increasing. If ω is bounded, this amounts to the constant case. We are thus led to assuming that ω is strictly increasing to infinity.

Of course the problem arises to determine whether a similar phenomenon occurs for algebraic numbers of degree greater than 2, or for the classical transcendental numbers. Generally speaking, I expect the classical numbers to behave like almost all numbers. For instance, when

$$\psi(q) = 1/q^{1+\epsilon}$$

as in Roth's theorem, it is known that there is only a finite number of solutions for the inequality $|qe - p| < \psi(q)$ (Mahler-Popken). Using the same method as Mahler, I can show that if α is rational and $\neq 0$, then e^{α} also satisfies this property with the above function ψ, or even a slightly better one, as in Mahler's proof (cf. Schneider [4]). For arbitrary ψ, the answer is not known, even in the case of algebraic numbers.

Finally, it should be mentioned that the brute force and essentially simple-minded methods of the present paper have of course no chance of succeeding when dealing with the approximation problem to algebraic numbers of higher degree (and even less to transcendental ones). I see no hints anywhere (in Roth's proof or elsewhere) for a method which would succeed in these more general cases. The combinatorial structure of all known proofs of measures of irrationality or transcendence never exhibits the appearance of the integral

$$\int_1^B \psi(t)\,dt$$

or the sum $\sum \psi(q)$. This seems to me to be the main reason why the technique of the Thue-Siegel-Schneider-Roth proof, involving a polynomial in several variables, fails to apply for numbers other than algebraic ones. For instance, instead of making certain coefficients equal to 0 in the proof, when dealing with values of say e^t at algebraic α, one could require only that such values are very small, by imposing a zero of high order on the functions $F^{(i)}(e^t, \cdots, e^t)$. However, the linear equations involved in this question have too many variables to allow one to solve them in the frame of the present structure of the proof.

COLUMBIA UNIVERSITY.

REFERENCES.

[1] E. Hecke, "Über analytische Funktionen und die Verteilung von Zahlen mod eins," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, Bd. 1 (1921), pp. 54-76.

[2] S. Lang, "Asymptotic approximation to quadratic irrationalities I," *American Journal of Mathematics*, vol. 87 (1965), pp. 481-487.

[3] S. Schanuel, "Heights in number fields," *to appear.*

[4] T. Schneider, *Einführung in die tranzoendente Zahlen*, Springer Verlag (1957), p. 88, and p. 91, where the argument with rational a works. When a is algebraic, the method gives only a weaker theorem, involving the degree of a.

STONE'S 2-SPHERE CONJECTURE.*

By F. Burton Jones.

In his definitive paper [8] of thirty years ago van Kampen pointed out the characterizations of the 2-sphere lean heavily on some form of Jordan curve theorem [6, 7, 10] or on some form of unicoherence of the Janiszewski-Mullikin type [5, 11]. Since then D. W. Hall [3] and R. H. Bing [1, 2] have considerably sharpened results of the first type to the point where those of Bing may very well be in final form. (Gail Young [9] has done a somewhat similar thing for van Kampen's characterization of a 2-manifold.) However, no such progress has been made on results of the second (or unicoherence) type. The purpose of this paper is to initiate this progress along the line suggested (in correspondence) by A. H. Stone.

Property (U). If the boundaries of two connected domains have a connected intersection, then the domains have a connected intersection.

Clearly property (U) holds true in an arc (but not in a simple closed curve [a connected set may be degenerate or empty]). Assuming that the space is separated by no point eliminates this trivial case. The property fails to hold for the 3-ball as can be seen by taking for the two domains the complements (in the 3-ball) of the right and left halves (respectively) of the equatorial disk. In fact, every cyclic Peano continuum in which property (U) holds true is a 2-sphere (Corollary 5).

Notation. Let S denote a connected, locally connected, compact metric space such that (1) no point separates S and (2) property (U) holds true in S.

Theorem 1. *In S no point separates a connected domain.*

Corollary 1. *No finite point set separates S (in particular, no pair of points separates S).*

Theorem 2. *If T is a closed point set which separates S, U_1 and U_2 are distinct complementary domains of T such that $\bar{U}_1 \supset T$ and $\bar{U}_2 \supset T$, then T is a continuum.*

---

* This work was supported by NSF GP-25.

Received August 6, 1964.

THEOREM 3. *No arc separates S.*

Proof. Suppose that some arc T separates the point x from the point y. Let aob denote a subarc of T irreducible with respect to being a closed subset of T which separates x from y. Let D_1 denote the complementary domain of the arc ao (of aob) which contains $x+y$ and let D_2 be the complementary domain of the arc ob (of aob) containing $x+y$. The intersection of D_1 and D_2 is not connected but the intersection of their boundaries is the point o. This contradicts property (U).

THEOREM 3.1. *If the simple closed curve C separates the point p from the point q (in S), then C is the boundary of each of its complementary domains containing p or q. Furthermore, if I is the complementary domain of C containing p, xay is an arc spanning C from $I-p$ and $C=xby+xcy$, then one of the two simple closed curves $xay+xby$ and $xay+xcy$ separates p from q.*

THEOREM 4. *If p is a point of a domain R, there exists a simple domain D such that $p \in D \subset R$ and $S-\bar{D}$ is a simple domain.*

Proof. Let U denote a connected domain such that $p \in U \subset \bar{U} \subset R$. Since no point separates U (Theorem 1), some continuum K lies in $R-p$ and contains the boundary of U. Now let G be a brick partitioning of S with a mesh small enough so that the closure of no element of G contains both a point of K and a point of $p+(S-R)$ [2]. Let M denote the sum of the closures of all elements of G whose closures contain p and let U_1 denote the interior of the point set consisting of M plus the closures of all elements of G which are separated from K by M. Clearly U_1 is a connected domain and $p \in U_1 \subset \bar{U}_1 \subset R$. Also if T denotes the boundary of U_1, then $S-(U_1+T)$ is a connected domain U_2 such that T is also the boundary of U_2. So by Theorem 2, T is a continuum.

To see that T is locally connected let p' be a point of T, let R' be a domain containing p', let U' be a connected domain such that $p' \in U' \subset \bar{U}' \subset R'$, and let K' be a continuum lying in $R'-p'$ which contains the boundary of U'. As in the first part of the proof of this theorem let G' be a brick partitioning of S which refines G and has mesh so small that the closure of no element of G' contains both a point of K' and a point of $p'+(S-R')$. Let M' denote the sum of the closures of all elements of G' whose closures contain p' and let U'_1 denote the interior of the point set consisting of M' plus the closures of all elements of G' which are separated from K' by M'. As above, U'_1 is a connected domain whose boundary B' is connected. Suppose

that $T \cdot U'_1$ is the sum of two mutually separated sets F_1 and F_2. Then $U'_1 - F_1$ and $U'_1 - F_2$ are connected domains and the intersection of their boundaries is B'. But $(U'_1 - F_1) \cap (U'_1 - F_2)$ is $U'_1 - T$ which is not connected. This is contrary to property (U). Hence $T \cdot U'_1$ is connected. It follows that T is locally connected. (One may use Theorem A, p. 242 of [7] but no material simplification results.)

Suppose that some subcontinuum N of T separates T, i.e.,

$$T - N = H + K, \qquad \bar{H} \cdot K = H \cdot \bar{K} = 0.$$

The two connected domains, $S - (H + N)$ and $S - (K + N)$, have a disconnected interesection but the intersection of their boundaries is the continuum N. This is contrary to property (U); so no subcontinuum of T separates T. It follows from one of Kuratowski's theorems that T is a simple closed curve.

LEMMA. *Suppose that the simple closed curve L separates (in S) the points p and q of the simple closed curve J from each other. Letting J_p and J_q denote the components of $J - J \cdot L$ containing p and q respectively, $[J - (J_p + J_q)] + L$ contains the boundary of a simple domain I such that I contains J_p but no point of $(J - J_p) + L$.*

Proof. Let I denote the component of $S - \{[J - (J_p + J_q)] + L\}$ which contains p. Clearly I contains no point of J_q. Since any subcontinuum of $[J - (J_p + J_q)] + L$ is locally connected, a slight variation of the argument to prove Theorem 4 shows that the boundary of I contains a simple closed curve C which separates p and q. Furthermore, if C were not all of the boundary of I, then Theorem 3.1 would be violated. So I is a simple domain.

THEOREM 5. *Every simple closed curve separates S.*

Proof. Suppose, on the contrary, that J is a simple closed curve such that $S - J$ is connected. Let p and q denote two distinct points of J and let D denote a simple domain containing p such that $S - \bar{D}$ is a simple domain containing q. The closure of the component of $J \cdot (S - \bar{D})$ containing q is an arc xqy of J. Let u and v denote points of the simple closed curve $\bar{D} - D$ which are separated in $\bar{D} - D$ by $x + y$ so that $\bar{D} - D$ is the sum of two arcs xuy and xvy. Let C_1 be the component of $S - (xqy + xuy)$ which contains p and let C_2 be the component or $S - (xqy + xvy)$ containing p. The boundary B_1 of C_1 is $xqy + xuy$, for otherwise some arc separates S contrary to Theorem 3. Likewise, the boundary B_2 of C_2 is $xqy + xvy$. Since

$B_1 \cdot B_2$ is xqy, it is connected and by property (U), $C_1 \cdot C_2$ is connected. But $C_1 \cdot C_2$ contains D; so $C_1 \cdot C_2 = D$. Since q is in the boundary of each of C_1 and C_2, each must intersect $S - \bar{D}$. Hence $(xqy) \cdot (S - \bar{D})$ separates $S - \bar{D}$. Each of the components of $(S - \bar{D}) - xqy$ has a simple closed curve for its boundary including xqy. The lemma shows that no loss of generality occurs if we assume that none of these components contains a point of J. Furthermore, each of these components has a limit point in $u + v$. Let uv denote an arc in $S - J$ from u to v. Repeating the argument for a smaller simple domain D' substituted for D (taking D' to be a small enough subset of D to miss uv) we get a contradiction because only one component of $(S - \bar{D}') - x'qy'$ will contain all of the components of $(S - \bar{D}) - xqy$ and hence only one component of $(S - \bar{D}') - x'qy'$ can have q as a limit point.

COROLLARY 5. *The space S is a 2-sphere* [1 *or* 10].

THEOREM 6. *Property* (U) *holds true in the 2-sphere S^2* [5].

Proof. Suppose that D_1 and D_2 are connected domains in S^2 whose boundaries B_1 and B_2 have a connected intersection T. Suppose that the points x and y belong respectively to different components U and V of $D_1 \cdot D_2$. Let B denote the outer boundary of U with respect to y. Since B is a continuum, every component of $B - B \cdot T$ has a limit point in T and those lying in B_1 together with T form a continuum T_1. Likewise those lying in B_2 together with T form a continuum T_2. Since T_1 and T_2 lie in B_1 and B_2 respectively, neither separates x from y. Since their intersection is connected their sum does not separate x from y. This is a contradiction.

THE UNIVERSITY OF CALIFORNIA, RIVERSIDE.

REFERENCES.

[1] R. H. Bing, "The Kline sphere characterization problem," *Bulletin of the American Mathematical Society*, vol. 52 (1945), pp. 644-653.

[2] ———, "Complementary domains of continuous curves," *Fundamenta Mathematicae*, vol. 36 (1949), pp. 303-318.

[3] D. W. Hall, "A partial solution of a problem of J. R. Kline," *Duke Mathematical Journal*, vol. 9 (1942), pp. 893-901.

[4] F. B. Jones, "On the existence of a small connected open set with a connected boundary," *Bulletin of the American Matematical Society*, vol. 68 (1962), pp. 117-119.

[5] C. Kuratowski, "Une caractérization de la surface de la sphére," *Fundamenta Mathematica*, vol. 13 (1929), pp. 307-318.

[6] R. L. Moore, "Concerning a set of postulates for plane analysis situs," *Transactions of the American Mathematical Society*, vol. 17 (1916), pp. 131-164.

[7] ———, *Foundations of point-set theory*, American Mathematical Society Colloquium Publications, vol. 13, Revised Edition. Providence, Rhode Island, 1962.

[8] E. R. van Kampen, "On some characterizations of 2-dimensional manifolds," *Duke Mathematical Journal*, vol. 1 (1935), pp. 74-93.

[9] G. S. Young, "A characterization of 2-manifolds," *Duke Mathematical Journal*, vol. 14 (1947), pp. 979-990.

[10] Leo Zippin, "On continuous curves and the Jordan curve theorem," *American Journal of Mathematics*, vol. 52 (1930), pp. 331-350.

[11] ———, "A study of continuous curves and their relation to the Janiszewski-Mullikin theorem," *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 744-770.

ON THE GRADED RING OF SIEGEL MODULAR FORMS OF GENUS TWO.*

By William F. Hammond.

Introduction. In this note we shall give a simple proof of the structure theorem for the graded ring of Siegel modular forms of even weight in the genus two case. The structure of this ring was determined by J. Igusa [3] several years ago. He found it to be the polynomial ring generated over the field of complex numbers by two modular forms of weights four and six and by two cusp forms of weights ten and twelve. More recently Igusa [5] has obtained a second proof from more general results.

The method which we shall employ here is the genus two analogue of a classical method for studying elliptic modular forms. The essence of the method is the construction of a restriction homomorphism on the ring under consideration (to a known ring) whose kernel is a principal ideal and whose image is computable. It is by this method that K.-B. Gundlach [1] has recently determined the structure of the graded ring of Hilbert modular forms for the field $\boldsymbol{Q}(\sqrt{5})$ (where $\boldsymbol{Q}$ denotes the field of rational numbers), and his work provided the motivation for ours.

In the elliptic case [cf. 2] this homomorphism is merely the restriction of a modular form to the point at infinity. The kernel is generated by a modular form of weight twelve whose only zeros are of order one at "cusps." In the genus two case we can conclude from Igusa's results that the cusp form of weight ten has similar properties. But in order to obtain a new proof of the structure theorem, it is necessary to prove these properties directly. This has been done by Igusa, and he was kind enough to offer us this part of his unpublished work for Section 2 of this paper.

1. Preliminaries. For the convenience of the reader we shall summarize some known material [cf. 4, 5, 6]. A *characteristic of genus* g is a pair $\mathfrak{m} = (m', m'')$ where m' and m'' are integer column vectors with g components. A characteristic is called *even* (or *odd*) when the scalar product

Received September 9, 1964.

* The author was supported in part as a research assistant under a grant from the National Science Foundation.

of m' and m'' is even (or odd). The *theta-constant of characteristic* $\mathfrak{m}$ is the function defined by

$$\theta_{\mathfrak{m}}(\tau) = \sum \exp \pi i\{{}^t(p+\tfrac{1}{2}m')\tau(p+\tfrac{1}{2}m') + {}^t(p+\tfrac{1}{2}m')m''\},$$

where the summation is taken over all integer column vectors p with g components, and where τ is a point of the Siegel upper-half plane of genus g, i.e., τ is a complex symmetric g-by-g matrix with positive definite imaginary part. The function $\theta_{\mathfrak{m}}$ is identically zero if and only if the characteristic $\mathfrak{m}$ is odd, and $\theta_{\mathfrak{m}}$ is determined up to a sign by $\mathfrak{m} \bmod 2$. Let θ denote the product of all (non-zero) $\theta_{\mathfrak{m}}$ taken over the finite set of even characteristics $\mathfrak{m}$ whose components are either 0 or 1. In the genus one case θ^8 is the cusp form of weight twelve. We shall see that θ^2 plays the same role in genus two as θ^8 in genus one.

The *Siegel modular group of genus* g is the matrix group consisting of all $2g$-by-$2g$ integer matrices M such that ${}^tMIM = I$ where I is the matrix

$$I = \begin{pmatrix} 0 & 1_g \\ -1_g & 0 \end{pmatrix}$$

with 1_g denoting the g-by-g unit matrix. The *principal congruence subgroup of level* q consists of those M which are congruent $\bmod\, q$ to the $2g$-by-$2g$ unit matrix. The operation of the modular group on the Siegel upper-half plane of genus g is given by $M\cdot\tau = (a\tau + b)(c\tau + d)^{-1}$ where a, b, c, d are the g-by-g components of M. If H is a subgroup of the modular group, we say that a holomorphic function f in the upper-half plane is a *Siegel modular form of weight* w *belonging to* H if we have

$$f(M\cdot\tau) = \det(c\tau + d)^w f(\tau)$$

for every M in H. When H is not otherwise designated and we speak of "modular forms," it is understood that H is the full modular group.

From this point we assume that $g = 2$. If we take the quotient group of the modular group by its principal congruence subgroup of level two, we obtain a group of order 720. This group has a faithful representation as a group of permutations of the set of six odd characteristics mod 2, and, therefore, it is isomorphic to the symmetric group on six letters. If M is a modular matrix, we let $\epsilon(M) = 1$ when the residue of M (mod 2) corresponds to an even permutation, and $\epsilon(M) = -1$ otherwise. We shall let Z denote the subgroup of index two in the modular group defined by $\epsilon(M) = 1$. A modular form f of weight w belonging to Z will be called *even* if it belongs to the full modular group and *odd* when for every modular matrix M we have

$$f(M\cdot\tau) = \epsilon(M)\det(c\tau + d)^w f(\tau).$$

The function θ is an odd modular form of weight five.

2. The zeros of θ. It is known [cf. 5] that θ has an expansion (in w_{12}) of the following form:

$$\theta\begin{pmatrix} w_1 & w_{12} \\ w_{12} & w_2 \end{pmatrix} = -2^7\pi i(\exp \pi i w_1)\,(\exp \pi i w_2)\,w_{12} + \cdots.$$

Hence θ vanishes to the first order along the diagonal plane $w_{12} = 0$. We need the following lemma in order to see that θ divides any even or odd modular form which vanishes on the diagonal plane:

LEMMA. *If τ is fixed, then $\theta(\tau) = 0$ if and only if there is a modular matrix M such that $M \cdot \tau = \tau'$ lies on the diagonal plane, i.e., we have*

$$\tau' = \begin{pmatrix} w_1 & 0 \\ 0 & w_2 \end{pmatrix}.$$

Proof. Since the "if" part is clear, we shall attack the "only if" part. We consider the complex torus A with $(\tau 1_2)$ as period matrix. It is well known that A is an abelian variety of dimension two. The sixteen theta-functions of characteristic $\mathfrak{m}$ defined by

$$\theta_{\mathfrak{m}}(\tau, z) = \sum_p \exp \pi i\{{}^t(p + \tfrac{1}{2}m')\tau(p + \tfrac{1}{2}m') + {}^t(p + \tfrac{1}{2}m')(2z + m'')\}$$

determine positive divisors on A. These positive divisors have self-intersection number two and are stable under the automorphism of A defined by substituting $-z$ for z. Moreover, they are translations of any one of them by points of ${}_2A$, i.e., points of order two on A. If we denote the divisor corresponding to $\mathfrak{m} = \mathfrak{o}$ by Θ, then the sixteen divisors are of the form Θ_s with s in ${}_2A$. Suppose that Θ_s corresponds to $\mathfrak{m}$. Then we have $\theta_{\mathfrak{m}}(\tau) = 0$ if and only if Θ_s contains the point o of A, i.e., if and only if Θ contains s. Therefore Θ certainly contains the six points of ${}_2A$ which correspond to odd characteristics. Now since we have $\theta(\tau) = 0$ by assumption, we have $\theta_{\mathfrak{m}}(\tau) = 0$ for some even $\mathfrak{m}$. Therefore Θ contains at least seven points of ${}_2A$. It has been observed by A. Weil [7, Theorem 2] that either (case 1) Θ is an irreducible non-singular curve of genus two, and the immersion $\Theta \to A$ defines the Jacobian variety of Θ; or (case 2) A is the product $A_1 \times A_2$ of two elliptic curves A_1 and A_2, and Θ is of the form $(s_1 \times A_2) + (A_1 \times s_2)$ where s_i is a point of ${}_2A_i$ for $i = 1, 2$. Now in case 1 the curve Θ can be considered as a two-sheeted covering of the rational curve which is the quotient of Θ by the automorphism $z \to -z$, and points of ${}_2A$ on Θ correspond precisely to the points of ramification of this covering. Since the genus of Θ is two, the number of points of ramification is six. Therefore the number of points of ${}_2A$ on Θ must be six. Since we know that Θ contains at least seven points

of ${}_2A$, we have case 2. Consequently, we can find a new coordinate system z' depending linearly on z, a point τ' of the form stated in the lemma, and a theta-function

$$\theta_{\text{II}}(\tau', z') = \theta_1(w_1, z_1')\theta_2(w_2, z_2')$$

such that $(w_i 1)$ is the period matrix of A_i, θ_i is the elliptic theta-function of characteristic (n_i', n_i''), and $\theta_i(w_i, z_i')$ determines the divisor s_i on A_i for $i = 1, 2$. We then have that $\theta_{\text{II}}(\tau', z')$ and $\theta_0(\tau, z)$ differ by the so-called trivial theta-function. As we know [cf. 4, p. 226], this implies that $\tau' = M \cdot \tau$ for some modular matrix M. This proves the lemma.

3. The structure theorem. Let us recall [cf. 6] that the *Eisenstein series of weight w* is the holomorphic function G_w defined in the Siegel upper-half plane by

$$G_w(\tau) = \sum \det(c\tau + d)^{-w}$$

where the summation is extended over the bottom halves (cd) of modular matrices M which vary over distinct cosets of the subgroup of the modular group defined by $c = 0$. We may formulate Igusa's structure theorem formally as follows:

THEOREM. *The graded ring of Siegel modular forms of even weight (and genus two) is a polynomial ring in four (algebraically independent) variables over the field of complex numbers. It is generated by the four Eisenstein series G_4, G_6, G_{10}, and G_{12}.*

Proof. If f is holomorphic in the Siegel upper-half plane, let $f^\#$ denote the restriction of f to the diagonal plane. By what we have said in the previous section, θ divides any even or odd modular form f for which $f^\# = 0$. Let S denote the graded ring of (even) modular forms of even weight. We consider the #-homomorphism restricted to S. If f is in S_w, i. e., if f is an element of S of weight w, then $f^\#$ is a holomorphic function of two variables which is an elliptic modular form in each variable when the other variable is fixed. If E denotes the graded ring of elliptic modular forms, then [cf. 8, p. 334] $f^\#$ is in $E_w \otimes E_w$, where the tensor product is taken with respect to the field of complex numbers. Moreover, we have $f^\#(t_1, t_2) = f^\#(t_2, t_1)$. Suppose that f is in S_w and $f^\# = 0$. Then $g = f/\theta$ is certainly an odd modular form. Under the substitution which changes the sign of the coefficients of τ which are not on the principal diagonal, f is invariant, while θ changes its sign. Hence g changes its sign, and thus $g^\# = 0$. So θ divides g, or θ^2 divides f. We have shown that the kernel of the #-homomorphism is the principal ideal in S generated by θ^2.

Fix w and let n denote the dimension of E_w. Then the dimension of $E_w \otimes E_w$ is n^2 and the dimension of its linear subspace defined by $f(t_1, t_2) = f(t_2, t_1)$ is $\frac{1}{2}n(n+1)$. In the appendix of [3] Igusa shows by an elementary argument that $G_4{}^{\#}$, $G_6{}^{\#}$, and $G_{12}{}^{\#}$ are algebraically independent over the complex numbers. Let R denote the graded subring of S generated by G_4, G_6, and G_{12}. Then the dimension of R_w is the number of non-negative integral solutions of the Diophantine equation $4p + 6q + 12r = w$. For fixed r the number of solutions is the dimension of E_{w-12r}. Observing that the dimension of E_w is one more than the dimension of E_{w-12}, and summing over r, we see that the dimension of R_w is $\frac{1}{2}n(n+1)$. Hence the image of R_w fills up the "symmetric subspace" of $E_w \otimes E_w$. We can now show S is generated by G_4, G_6, G_{12}, and θ^2. Let f be an arbitrary element of S_w. If we choose h in R_w such that $f^{\#} = h^{\#}$, then there is an element g of S_{w-10} such that $f - h = \theta^2 g$. By induction we may assume that g is in the subring of S generated by the three Eisenstein series and θ^2. Hence f is also in this subring. Since the dimension of E_{10} is one, it is clear that $G_4{}^{\#}G_6{}^{\#} = G_{10}{}^{\#}$. After looking at some Fourier coefficients [cf. 3] to ensure that $G_4G_6 - G_{10}$ is not identically zero, we find that $G_4G_6 - G_{10}$ differs by a constant factor from θ^2. This completes the proof of the structure theorem.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] K. B. Gundlach, "Die Bestimmung der Funktionen zur Hilbertschen Modulgruppe des Zahlkörpers $Q(\sqrt{5})$," *Mathematische Annalen*, vol. 152 (1963), pp. 226-256.

[2] E. Hecke, "Die Primzahlen in der Theorie der Elliptischen Modulfunktionen," *Mathematische Werke*, Göttingen, 1959, pp. 577-590.

[3] J. Igusa, "On Siegel modular forms of genus two," *American Journal of Mathematics*, vol. 84 (1962), pp. 175-200.

[4] ———, "On the graded ring of theta-constants," *American Journal of Mathematics*, vol. 86 (1964), pp. 219-246.

[5] ———, "On Siegel modular forms of genus two (II)," *American Journal of Mathematics*, vol. 86 (1964), pp. 392-412.

[6] C. L. Siegel, "Einführung in die Theorie der Modulfunktionen n-ten grades," *Mathematische Annalen*, vol. 116 (1939), pp. 617-657.

[7] A. Weil, "Zum Beweis des Torellischen Satzes," *Göttingen Nachrichten*, Nr. 2 (1957), pp. 33-53.

[8] E. Witt, "Eine Identität zwischen Modulformen zweiten Grades," *Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität*, vol. 14 (1941), pp. 323-337.

STUDIES IN EQUISINGUARITY I
EQUIVALENT SINGUARITIES OF PLANE ALGEBROID CURVES.

By Oscar Zariski.*

With this paper we initiate a series of investigations on the concept of equisingularity of an algebraic variety V, along an irreducible (singular) subvariety W of V, at a given point of W. The actual definition of this concept, at least in certain special situations (and, in particular, in the case $\operatorname{cod}_V W = 1$), was given by us in an earlier paper ("Equisingular points on algebraic varieties," seminari dell'Istituto Nazionale di Alta Matematics 1962-63, Roma, 1964, pp. 164-177), and will be presented and discussed in full (i. e., including proofs) in our next paper in this series. The present paper is preliminary in nature. It deals first of all with various ways of defining equivalent singularities of plane algebroid curves (§§ 1-4 bis). We define equivalence by induction on the number of locally quadratic transformations which are necessary to resolve the given singularity. No explicit (or implicit) use of Puiseux expansions is made, and, in fact, the case of non-zero characteristic (where Puiseux expansions are not available) is included in our treatment. We give 3 different definitions of equivalence, and we prove that they are all equivalent to each other. [It is our belief, however, that in the case of non-zero characteristic our treatment is not definitive, as it lumps together, in one equivalence class, singularities which, on the basis of some internal evidence, should be further subclassified; we propose to come back to the case $p \neq 0$ in a subsequent paper in this series.]

In § 5 we establish some technical results on derivations in complete integrally closed, domains (Theorems 3 and 4) and we obtain, as a consequence, a criterion for such a domain to be a power series ring over a subring.

In §§ 6-7, the geometric results of §§ 1-4 bis and the purely algebraic results of § 5 are applied toward the derivation of a discriminant criterion which enables one to decide whether or not, given an analytic family of algebroid curves C_t: $f(x, y; t) = 0$, the special member C_0 of the family has a singularity which is equivalent to the singularity of the generic member C_t.

Received October 26, 1964.

* This research was supported in part by the Air Force Office of Scientific Research and in part by the National Science Foundation.

This criterion will be used in our next paper dealing with equisingularity in codimension 1.

1. **Algebroid plane curves.** Let k be an algebraically closed field (of arbitrary characteristic) and let $k[[x,y]]$ be the ring of power series of two independent variables x, y, with coefficients in k. By an *algebroid plane curve* C we mean a local ring of the form $k[[x,y]]/(f)$, where $f=f(x,y)$ is a non-unit element of $k[[x,y]]$, free from multiple factors. We shall say that $f(x,y)=0$ is an *equation* of the curve C (the power series f is determined by C uniquely, to within an arbitrary unit factor in $k[[x,y]]$). The point $P: x=y=0$ will be referred to as the origin of C.

If $f(x,y)$ is an irreducible element of $k[[x,y]]$, C is said to be an irreducible algebroid curve. If f is reducible, $f=\phi_1\phi_2\cdots\phi_h$, where each ϕ_i is an irreducible element, then the h irreducible algebroid curves γ_i: $\phi_i=0$ are the irreducible *components* or *branches* of C.

Assume that the power series f begins with terms of degree s: $f=f_s(x,y)$ + terms of degree $>s$, where f_s is a homogeneous polynomial of degree s ($f_s=$ leading form of f). Then one says that P is an *s-fold point of C*, and we write $s=m_P(C)$. Clearly $m_P(C)=\sum_{i=1}^{h} m_P(\gamma_i)$.

It is known that if f is irreducible then f_s is the power of a linear form: $f_s=(ax+by)^s$ (Hensel's lemma). The line (or direction) $ax+by=0$ is then called the tangent line of C. If C is reducible, then the tangent lines of the various irreducible components γ_i of C shall be, by definition, the tangent lines of C (the number of distinct tangent lines of C is therefore $\leq h$).

2. **The quadratic transform of C.** Let l_i: $a_ix+b_iy=0$ be the tangent line of γ_i $(i=1,2,\cdots,h)$. Without loss of generality we may assume that the b_i are all different from zero. We set $x'=x$, $y'=y/x$, and

$$f(x',x'y')=x'^sf'(x',y').$$

Then f' is a power series in x', y', and in fact, f' is a power series in x', whose coefficients are *polynomials* in y' (with coefficients in k):

$$f'(x',y')=f_s(1,y')+x'f_{s+1}(1,y')+x'^2f_{s+2}(1,y')+\cdots \tag{1}$$

Let t be the number of *distinct* tangent lines p_ν of C, and let $y-\alpha_\nu x=0$ $(\nu=1,2,\cdots,t)$ be the equations of these lines (so that for each $i=1,2,\cdots,h$ the ratio $-a_i/b_i$ is equal to one of the α_ν). Let I_ν be the set of integers i $(1\leq i\leq h)$ such that $-a_i/b_i=\alpha_\nu$, and let C_ν be the union of the irreducible branches γ_i, $i\in I_\nu$. If we set $F_\nu(x,y)=\prod_{i\in I_\nu}\phi_i(x,y)$, then C_ν is the algebroid

curve defined by the equation $F_\nu(x,y)=0$. The irreducible branches of C_ν are the γ_i, with $i\in I_\nu$, and they all have the same tangent line p_ν. We have $f(x,y)=\prod_{\nu=1}^{t}F_\nu(x,y)$, so that C is the union of the t algebroid curves C_ν. We call the C_ν the *tangential components* of C.

Let $r_i=m_P(\gamma_i),s_\nu=m_P(C_\nu)\,(=\sum_{i\in I_\nu}r_i)$,

$$\phi_i(x,y)=\phi_{ir_i}(x,y)+\phi_{i,r_i+1}(x,y)+\cdots,$$
$$F_\nu(x,y)=F_{\nu,s_\nu}(x,y)+F_{\nu,s_\nu+1}(x,y)+\cdots.$$

Since $\phi_{ir_i}(x,y)=(y-\alpha_\nu x)^{r_i}$ for $i\in I_\nu$ and $F_{\nu,s_\nu}(x,y)=(y-\alpha_\nu x)^{s_\nu}$, it follows that

$$\phi_i(x',x'y')=x'^{r_i}\phi_i'(x',y'),$$
$$F_\nu(x',x'y')=x'^{s_\nu}F_\nu'(x',y'),$$

where, for $i\in I_\nu$:

$$\phi_i'(x',y')=(y'-\alpha_\nu)^{r_i}+x'\phi_{i,r_i+1}(1,y')+\cdots,$$
$$F_\nu'(x',y')=\prod_{i\in I_\nu}\phi_i'(x',y')=(y'-\alpha_\nu)^{s_\nu}+x'F_{\nu,s_\nu+1}(1,y')+\cdots.$$

We regard ϕ_i' and F_ν' as power series in x' and $y'-\alpha_\nu$ and we denote by γ_i' and C_ν' the algebroid curves, with origin at the point $P_\nu'(0,\alpha_\nu)$, defined respectively by the equations $\phi_i'(x',y')=0$ and $F_\nu'(x',y')=0$. The curves γ_i' $(i\in I_\nu)$ are irreducible branches of C_ν'. We note that

$$f'(x',y')=\prod_{\nu=1}^{t}F_\nu'(x',y').$$

Hence the equation $f'(x',y')=0$ may be regarded as representing the set of (disjoint) algebroid curves $C_1',C_2',\cdots,C_t'$, having respectively as origins the points $P_1',P_2',\cdots,P_t'$. We denote by T the locally quadratic transformation (with center P) defined by $x'=x$, $y'=y/x$, we set $C'=\bigcup_{\nu=1}^{t}C_\nu'$, and we call C', C_ν' and γ_i' the *proper quadratic transforms* of C, C_ν and γ_i respectively:

$$C'=T(C),$$
$$C_\nu'=T(C_\nu),$$
$$\gamma_i'=T(\gamma_i).$$

We shall refer to $C_1',C_2',\cdots,C_t'$ are the *connected components* of C'.

Note. Although the above definitions are based on a specific (and explicit) choice of regular parameters x, y of the point P of the (x,y)-plane, they are easily seen to have an intrinsic invariantive meaning, as they could

have been formulated in terms of local algebra (see Zariski-Samuel, Commutative Algebra, v. 2, Appendix 5). In such an intrinsic formulation one starts with a complete, noetherian, equicharacteristic local ring $\mathfrak{o}$, of dimension 1, whose residue field is algebraically closed and whose maximal ideal $\mathfrak{m}$ has a basis of two elements. If one, furthermore, assumes that $\mathfrak{o}$ has no proper nilpotent elements, then—for any given choice of a field k of representatives in $\mathfrak{o}$ and of a basis (ξ, η) of $\mathfrak{m}$—the local ring $\mathfrak{o}$ will be the local ring of a unique plane algebroid curve, defined over k. *The notion of the quadratic transform of* $\mathfrak{o}$ *is, however, independent of the choice of* k, ξ *and* η.

Similarly, given two local rings $\mathfrak{o}$, $\bar{\mathfrak{o}}$, satisfying the above conditions, and having the same (or isomorphic) residue fields, *then the definition of equivalence of* $\mathfrak{o}$ *and* $\bar{\mathfrak{o}}$ *given in* § 3 *below, remain meaningful and are independent of the choice of* k, ξ *and* η.

3. Three definitions of equivalence of algebroid curves. Let D be another plane algebroid curve, with some origin Q. We *assume that* C *and* D *have the same number* h *of irreducible branches*, and we denote by $\delta_1, \delta_2, \cdots, \delta_h$ the irreducible branches of D.

DEFINITION 1. *A* $(1,1)$ *mapping* π *of the set of branches* $\gamma_1, \gamma_2, \cdots, \gamma_h$ *of* C *onto the set of branches* $\delta_1, \delta_2, \cdots, \delta_h$ *of* D *is said to be a tangentially stable pairing* $\pi\colon C \rightarrow D$ *between the branches of* C *and those of* D, *if the following condition is satisfied: given any two branches* γ_i *and* γ_j *of* C, *the corresponding branches* $\pi(\gamma_i)$ *and* $\pi(\gamma_j)$ *of* D *have the same tangent if and only if* γ_i *and* γ_j *have the same tangent.*

Assume that there exists a tangentially stable pairing $\pi\colon C \rightarrow D$ between the branches of C and the branches of D. Then it is clear that C and D have the same number t of distinct tangent lines and that π induces a $(1,1)$ mapping of the set $\{p_1, p_2, \cdots, p_t\}$ of tangent lines of C onto the set $\{q_1, q_2, \cdots, p_t\}$ of tangent lines of D. We choose our indexing of these tangent lines in such a way that p_ν and q_ν are paired in this induced mapping, and we denote by C_ν (resp. D_ν) the tangential component of C (resp., D) associated with p_ν (resp., q_ν). Then it is clear that for each $\nu = 1, 2, \cdots, t$, π induces a $(1,1)$ mapping $\pi_\nu\colon C_\nu \rightarrow D_\nu$ of the set of branches of C_ν onto the set of branches of D_ν (the pairing π_ν is trivially tangentially stable, since both C_π and D_ν have only one tangent line).

Let π and π_ν be as above (π-tangentially stable), let T be a locally quadratic transformation with center at the origin P of C and let S be a locally quadratic transformation with center at the origin Q of D. Let $C' = T(C)$, $C_\nu' = T(C_\nu)$, $D' = S(D)$, $D_\nu' = S(D_\nu)$ be the proper transforms.

It is clear that π_ν induces a $(1,1)$ mapping π_ν' of the set of branches of C_ν' onto the set of branches of D_ν'. Namely, if we assume that the branches of C and D have been so indexed that $\pi(\gamma_i) = \delta_i$, for $i = 1, 2, \cdots, h$, then, in the notations of § 2, we set $\pi_\nu'(\gamma_i') = \delta_i'$ for all $i \in I_\nu$, where $\gamma_i' = T(\gamma_i)$ and $\delta_i' = S(\delta_i)$. The pairing $\pi_\nu': C_\nu' \to D_\nu'$ between the branches of C_ν' and the branches of D_ν' is, however, not necessarily tangentially stable.

An algebroid curve C is *regular* if its origin P is a simple point of C, i.e., if $m_P(C) = 1$. If P is a *singular* point (i.e., if $m_P(C) > 1$), then we can resolve the singularity of C at P by a finite number of locally quadratic transformations. By a sequence of *successive quadratic transforms of* C we mean a sequence $\{C, C', C'', \cdots, C^{(i)}, \cdots\}$ of algebroid curves $C^{(i)}$ such that for each i, $C^{(i+1)}$ is a connected component of the proper transform of $C^{(i)}$ under a locally quadratic transformation whose center is the origin of $C^{(i)} (C^{(0)} = C)$. The fact that the singularity of C can be resolved can then be stated as follows: there exists an integer N such that in *any* sequence of successive quadratic transforms of C, the curves $C^{(i)}$ are regular if $i \geqq N$. We denote $\sigma(C)$ the smallest integer N with the above property ($\sigma(C) = 0$ if and only if C itself is a regular curve).

It is clear that if $C_1', C_2', \cdots, C_t'$ are the connected components of the proper quadratic transform $T(C)$ of C, and if $\sigma(C) > 0$, then $\sigma(C_\nu') < \sigma(C)$ for $\nu = 1, 2, \cdots, t$. Our first definition of equivalence of algebroid curves (or—what is the same—of equivalence of algebroid singularities) proceeds by induction on $\sigma(C)$.

Let $\pi: C \to D$ be a pairing between the branches of C and the branches of D (it is already assumed that C and D have the same number h of branches). If C is regular (whence $\sigma(C) = 0$), then C (and therefore also D) has only one branch, $\pi: C \to D$ is uniquely determined, and we say that π is an (a)-*equivalence* if also D is a regular curve. Assume that for all pairs of algebroid curves Γ, Δ with the same number of branches and such that $\sigma(\Gamma) < \sigma(C)$ it has already been defined what is to be meant by saying that a pairing $\Gamma \to \Delta$ between the branches of Γ and the branches of Δ is an (a)-equivalence. Then we define an (a)-equivalence between C and D as follows (we use the notations introduced earlier in this section):

DEFINITION 2. *An (a)-equivalence $\pi: C \to D$ is a pairing π between the branches of C and the branches of D having the following properties*:

1) *π is tangentially stable.*

2) *If $\delta_i = \pi(\gamma_i)$ $(i = 1, 2, \cdots, h)$, then $m_P(\gamma_i) = m_P(\delta_i)$.*

3) *The pairing $\pi_\nu': C_\nu' \to D_\nu'$ $(\nu = 1, 2, \cdots, t)$ is an (a)-equivalence.*

Example. P is said to be an *ordinary* s-fold point of C if $s = m_P(C)$ and if the number t of distinct tangents of C is exactly equal to s. It follows that if P is an ordinary s-fold point of C, then C has exactly s branches $\gamma_1, \gamma_2, \cdots, \gamma_s$, these branches have distinct tangents $p_1, p_2, \cdots, p_s$, and each γ_i is a regular algebroid curve (since $s = \sum_{i=1}^{s} m_P(\gamma_i)$ and since $m_P(\gamma_i) \geqq 1$). The proper quadratic transform $C' = T(C)$ of C has then exactly s connected components $C_1', C_2', \cdots, C_s'$, where $C_i' = \gamma_i' = T(\gamma_i)$, and each C_i' is a regular curve (since it follows, quite generally, from (1), § 2, that $\sum_{\nu=1}^{t} m_{P_\nu'}(C_\nu') \leqq m_P(C)$, for any algebroid curve C). Now, suppose that P is an ordinary s-fold point of C and that a pairing $\pi\colon C \to D$ between the branches of C and D satisfies conditions 1) and 2) of Definition 2. Then it is clear that also Q is an ordinary s-fold point of D. Condition 3) is now automatically satisfied, since C_ν' and D_ν' are regular curves. Then π is an (a)-equivalence. Conversely, it is clear that if P is an ordinary s-fold point of C and if Q is an ordinary s-fold point of D, then any pairing $\pi\colon C \to D$ between the s branches of C and the s branches of D is an (a)-equivalence.

We now proceed to our second definition of equivalence between algebroid singularities. If T is our quadratic transformation, with center P (see § 2), then T blows up P into the line $x' = 0$ of the (x', y')-plane. We denote this line by $\mathcal{E}'$ and we refer to $\mathcal{E}'$ as the *exceptional curve* of T. If C_ν is a tangential component of C and $C_\nu' = T(C_\nu)$ is the proper T-transform of C_ν, then $\mathcal{E}'$ contains the origin P_ν' of C_ν', but $\mathcal{E}'$ is not a component of C_ν'. We denote by $C_\nu'^*$ the algebroid curve $C_\nu' \cup \mathcal{E}'$ and we call $C_\nu'^*$ the *total* T-transform of C_ν; in symbols: $C_\nu'^* = T\{C_\nu\}$. We set $C'^* = T(C) \cup \mathcal{E}'$ and we call C^* the *total* T-transform of C. Note that $m_{P_\nu'}(C_\nu'^*)$ is always $\geqq 2$.

It is known that after a finite number of successive quadratic transformations one can reach a stage where the total transform of C *has only ordinary double points.* More precisely: there exists an integer $N \geqq 0$ (depending on C) with the following property: if $\{C, C'^*, C''^*, \cdots, C^{(i)*}, \cdots\}$ is *any* sequence of algebroid curves such that for any i we have $C^{(i+1)*} = C^{(i+1)} \cup \mathcal{E}^{i+1}$, where $C^{(i+1)}$ is a connected component of the proper quadratic transform $T^{(i)}(C^{(i)*})$ of $C^{(i)*}$ ($T^{(i)}$ being a quadratic transformation with center at the origin $P^{(i)}$ of $C^{(i)*}$) and $\mathcal{E}^{(i+1)}$ is the exceptional curve of $T^{(i)}$, then for $i \geqq N$ the origin $P^{(i)}$ of $C^{(i)*}$ is an ordinary double point of $C^{(i)*}$. We denote by $\sigma^*(C)$ the smallest integer N having the above property.

It is clear that $\sigma^*(C) = 0$ if and only if the origin P of C is an ordinary double point of C. If C is a regular curve then a strict interpretation of our

definition of $\sigma^*(C)$ would require to set $\sigma^*(C) = 1$. However, we agree to set $\sigma^*(C) = 0$ also if C is a regular curve (this could also have been achieved by a slight change in our general definition of $\sigma^*(C)$). It is easily seen that $\sigma^*(C) = 1$ if and only if P is an ordinary s-fold point of C and $s > 2$. If P is a tacnode of the first kind (i.e., if C consists of two regular branches having simple contact), then the *total* transform $C^* = T\{C\}$ of C has an ordinary triple point, and hence $\sigma^*(C) = 2$ in this case. On the other hand, if P is an ordinary cusp of C, then $C'^* = T\{C\}$ has a tacnode of the first kind (while the *proper* transform $T(C)$ of C is regular), and thus $\sigma^*(C) = 3$ in this case.

Let C and D have the same number of branches and let $\pi :: C \to D$ be a pairing of the branches of C with the branches of D. If $\sigma^*(C) = 0$, i.e., if P is either a simple point or an ordinary double point of C, then we shall say that π is a *(b)-equivalence between C and D* if and only if also $\sigma^*(D) = 0$, i.e., if and only if the origin Q of D is a simple point or an ordinary double point of D according as P is a simple point or an ordinary double point of C. Assume that for all pairs Γ, Δ of algebroid curves, with the same number of branches, such that $\sigma^*(\Gamma) < \sigma^*(C)$, it has already been defined what is meant by saying that a pairing $\Gamma \to \Delta$ between the branches of Γ and the branches of Δ is a (b)-equivalence. Then we define a (b)-equivalence between C and D as follows:

DEFINITION 3. *A (b)-equivalence $\pi: C \to D$ is a pairing π between the branches of C and the branches of D having the following properties:*

1) *π is tangentially stable.*
2) *The pairings $\pi_\nu': C_\nu' \to D_\nu'$ $(\nu = 1, 2, \cdots, t)$ are (b)-equivalences.*
3) *If $\mathcal{E}'$ and E' are the exceptional curves of the quadratic transformations T and S respectively (having centers at P and Q), if $C_\nu'^* = C_\nu' \cup \mathcal{E}'$, $D_\nu'^* = D_\nu' \cup E'$, and if we extend the pairing π_ν' to a pairing $\pi_\nu'^*: C_\nu'^* \to D_\nu'^*$ by setting $\pi_\nu'^*(\mathcal{E}') = E'$, then $\pi_\nu'^*$ is a (b)*-equivalence.

Note that conditions 1) and 2) of this definition are identical with the condtions 1) and 3) of Definition 2; condition 2) of Definition 2 has been deleted and has been replaced in Definition 3 by condition 3). Thus the equality of the multiplicities of corresponding branches under π is not explicitly postulated in Definition 3.

We now give a third definition of equivalence of algebroid singularities, which we shall refer to as *formal equivalence.* Again we proceed by induction on $\sigma^*(C)$, where we agree that if $\sigma^*(C) = 0$ formal equivalence coincides with (b)-equivalence.

DEFINITION 4. *Given two algebroid curves C, D, having the same number of branches, we say that C and D are formally equivalent if there exists a tangentially stable pairing $\pi: C \to D$ between the branches of C and the branches of D such that (in our previous notations):*

1) *C_ν' and D_ν' are formally equivalent ($\nu = 1, 2, \cdots, t$).*
2) *$C_\nu'^*$ and $D_\nu'^*$ are formally equivalent ($\nu = 1, 2, \cdots, t$).*

Note that this definition does not say anything about the nature of the pairings $\pi_\nu': C_\nu' \to D_\nu'$ and $\pi_\nu'^*: C_\nu'^* \to D_\nu'^*$ induced by π. Condition 1) merely requires that there exist, for each $\nu = 1, 2, \cdots, t$, some tangentially stable pairing $\rho_\nu': C_\nu' \to D_\nu'$ satisfying the conditions of the above inductive definition; and similarly, condition 2) requires that there exists a tangentially stable pairing $\rho_\nu'^*: C_\nu'^* \to D_\nu'^*$ satisfying similar conditions. It is not even required that $\rho_\nu'^*$ be an extension of ρ_ν'. For this reason, Definition 4 is the most subtle (and also the weakest) of our three definitions of equivalence. The fact (established below) that these three definitions are all equivalent to each other is therefore not devoid of interest.

W shall use the following notations:

If $\pi: C \to D$ is a pairing between the branches of C and the branches of D, then we write $\pi: C \overset{a}{\underset{=}{\Longrightarrow}} D$ or $\pi: C \overset{b}{\underset{=}{\Longrightarrow}} D$ according as π is an (a)-equivalence of a (b)-equivalence. We shall write $C \overset{f}{\equiv} D$ if C and D are formally equivalent.

4. Proof of equivalence of Definitions 2, 3 and 4. We recall first a few (well-known) elementary facts about the intersection number (C, D) of algebroid curves (having the same origin P) and about the behaviour of this number under a quadratic transformation T centered at P. We assume, of course, that C and D have no branches in common.

If $n = m_P(C)$ and $m = m_P(D)$ and if we assume that the line $x = 0$ is not tangent to either C or D, then, by the Weierstrass properation theorem, we may assume that C and D are defined by $f(x, y) = 0$ and $g(x, y) = 0$, where f and g *are monic polynomials in* y, of degree n and m respectively, with coefficients which are power series in x. If $y_1, y_2, \cdots, y_n$ are the roots of $f(x, y)$ and $\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_m$ are the roots of $g(x, y)$ (in an algebraic closure of the field $k\{\{x\}\}$ of meromorphic functions of x), then (C, D), is defined as the order of the power series $\prod_{i=1}^{n} \prod_{j=1}^{m} (y_i - \bar{y}_j)$ in x (this power series is an element of $k[[x]]$ since the y_i and $\bar{y}_j$ are integral over $k[[x]]$). Since n and m are the degrees of the leading forms of f and g, it follows that

$$f = y^n + xa_1(x)y^{n-1} + \cdots + x^n a_n$$
$$g = y^m + xb_1(x)y^{m-1} + \cdots + x^m b_m(x),$$

with $a_i(x)$ and $b_j(x)$ in $k[[x]]$. If $f'(x',y') = 0$ and $g'(x',y') = 0$ are the equations of the proper transforms $C' = T(C)$ and $D' = T(D)$ of C and D (see § 2), then

$$f' = y'^n + a_1(x')y'^{n-1} + \cdots + a_n(x')$$
$$g' = y'^m + b_1(x')y'^{m-1} + \cdots + b_m(x'),$$

where $x' = x$. Thus, if $y_1', y_2', \cdots, y_n'$ and $\bar{y}_1', \bar{y}_2', \cdots, \bar{y}_m'$ are the roots of $f'(x',y')$ and $g'(x',y')$ (regarded as polynomials in y'), then the y_i' and $\bar{y}_j'$ are still integrally dependent on $k[[x]]$. Since $y_i = xy_i'$ and $\bar{y}_j = x\bar{y}_j'$, it follows that if $C_1', C_2', \cdots, C_t'$ are the connected components *of* C' and $D_1', D_2', \cdots, D_\tau'$ are the connected components of D', then

$$(2) \qquad (C,D) = mn + \sum_{\nu=1}^{t} \sum_{\mu=1}^{\tau} ((C_\nu', D_\mu').$$

Here $(C_\nu', D_\mu') > 0$ if and only if C_ν' and D_μ' have the same origin. By (2) if follows that $(C,D) \geqq mn$, with equality if and only if C and D have no common tangents. From the above expression of f' it follows also that if $\mathcal{E}' : x' = 0$ is the exceptional curve of T, then

$$(3) \qquad (C', \mathcal{E}') = \sum_{\nu=1}^{t} (C_\nu', \mathcal{E}') = n (= m_P(C)).$$

These are all the facts that we shall need in the sequel.

Let now C and D be algebroid curves, with origins P and Q, and with the same number of branches.

LEMMA 1. *Let* $\pi : C \overset{a}{\Longrightarrow} D$ *be an* (a)*-equivalence between* C *and* D *and let* Γ *be a regular curve through* P *and* Δ *a regular curve through* Q. *Assume that* Γ *is not a branch of* C, *that* Δ *is not a branch of* D, *and that for every pair* (γ_i, δ_i) *of corresponding branches under* π *we have*

$$(4) \qquad (\gamma_i, \Gamma) = (\delta_i, \Delta), \qquad i = 1, 2, \cdots, h.$$

Then the pairing $\rho : C \cup \Gamma \to D \cup \Delta$ *between the branches of* $C \cup \Gamma$ *and the branches of* $D \cup \Delta$ *which is the unique extension of* π *by* $\rho(\Gamma) = \Delta$, *is an* (a)*-equivalence.*

Proof. (by induction on $\sigma^*(C \cup \Gamma)$).

If $\sigma^*(C \cup \Gamma) = 0$, then C is a regular curve, and the regular curves C and Γ have distinct tangents. Hence, by (2), we have $(C, \Gamma) = 1$. Since

C and D are (a)-equivalent, also D is a regular curve, and it follows by (4) that $(D, \Delta) = 1$, whence $\sigma^*(D \cup \Delta) = 0$. Therefore ρ is an (a)-equivalence.

In the general case, we observe that $(\gamma_i, \Gamma) \geqq m_P(\gamma_i)$, with equality if and only if γ_i and Γ do not have the same tangent. Similarly, for (δ_i, Δ) and $m_Q(\delta_i)$. Since π is an (a)-equivalence, we have $m_P(\gamma_i) = m_Q(\delta_i)$. Hence it follows from (4) that γ_i and Γ have the same tangent if and only if δ_i and Δ have the same tangent. Therefore *ρ is a tangentially stable pairing.* Since $m_P(\Gamma) = m_Q(\Delta) = 1$, it is clear that any two corresponding branches under ρ have the same multiplicity. Thus ρ satisfies conditions 1) and 2) of the definition of an (a)-equivlance (Definition 2, § 3). We now check condition 3) of that definition.

Let $T(\Gamma) = \Gamma'$, $S(\Delta) = \Delta'$. We consider two cases.

First Case. Γ is not tangent to any of the branches of C (and hence Δ is not tangent to any of the branches of D). In this case, $C_1', C_2', \cdots, C_t'$ and Γ' are the connected components of the proper transform $T(C \cup \Gamma)$, and $D_1', D_2', \cdots, D_t'$ and Δ' are the connected components of $S(D \cup \Delta)$. If $\rho_\nu' \colon C_\nu' \to D_\nu'$ $(\nu = 1, 2, \cdots, t)$, $\rho_{t+1}' \colon \Gamma' \to \Delta'$ are the pairings induced by ρ, then $\rho_\nu' = \pi_\nu'$ is an (a)-equivalence (since π is an (a)-equivalence), and ρ_{t+1}' is trivially an (a)-equivalence (since Γ' and Δ' are regular curves). Thus condition 3) of Definition 2 is satisfied.

Second Case. The tangent of Γ coincides with one of the t tangents $p_1, p_2, \cdots, p_t$ of C, say p_1 is the tangent of Γ. In that case, the corresponding tangent line q_1 of D (under π) is the tangent of Δ. The tangential components of $C \cup \Gamma$ are now $C_1 \cup \Gamma$, $C_2, \cdots, C_t$, and the corresponding tangential components (under ρ) of $D \cup \Delta$ are $D_1 \cup \Delta, D_2, \cdots, D_t$. The connected components of $T(C \cup \Gamma)$ and $S(D \cup \Delta)$ are now $C_1' \cup \Gamma'$, $C_2', \cdots, C_t'$ and $D_1' \cup \Delta'$, $D_2', \cdots, D_t'$. Consider the induced pairing

$$\rho_1' \colon C_1' \cup \Gamma' \to D_1' \cup \Delta', \quad \rho_\nu' \colon C_\nu' \to D_\nu' \qquad (\nu = 2, 3, \cdots, t).$$

We have $\rho_\nu' = \pi_\nu'$ for $\nu \geqq 2$, and ρ_ν' is an (a)-equivalence for $\nu = 2, 3, \cdots, t$. Consider now ρ_1'. It is an extension of π_1', with $\rho_1'(\Gamma') = \Delta'$. Let γ_i' be any branch of C_1' $(i \in I_1)$ and let δ_i' be the corresponding branch of D_1' (under π_1'). Then $\gamma_i' = T(\gamma_i)$, $\delta_i' = S(\delta_i)$, with $\delta_i = \pi(\gamma_i)$. By (2), we have $(\gamma_i, \Gamma) = m_P(\gamma_i) + (\gamma_i', \Gamma')$, $(\delta_i, \Delta) = m_P(\delta_i) + (\delta_i', \Delta')$. Since $m_P(\gamma_i) = m_P(\delta_i)$ and $(\gamma_i, \Gamma) = (\delta_i, \Delta)$, it follows that $(\gamma_i', \Gamma') = (\delta_i', \Delta')$. Since $\sigma^*(C_1' \cup \Gamma') < \sigma(C \cup \Gamma)$ (if $\sigma^*(C \cup \Gamma) \neq 0$), it follows, by our induction hypothesis, that ρ_1' is an (a)-equivalence. This completes the proof of the lemma.

LEMMA 2. *If* $\pi: C \overset{a}{\Longrightarrow} D$ *and* $\pi(\gamma_i) = \delta_i$ $(i = 1, 2, \cdots, h)$, *then*

$$\gamma_i, \gamma_j) = (\delta_i, \delta_j) \text{ for all } i \neq j.$$

Proof. We have $(\gamma_i, \gamma_i) = m_P(\gamma_i) m_P(\gamma_j) + (\gamma_i', \gamma_j')$, where $\gamma_i' = T(\gamma_i)$, $(\delta_i, {}_j) = m_Q(\delta_i) m_Q(\delta_j) + (\delta_i', \delta_j')$. Here (γ_i', γ_j') is to be replaced by zero if γ_i' and γ_j' have distinct origins (and similarly for (δ_i', δ_j')). Since $m_P(\gamma_i) = m_Q(\delta_i)$ for all i, the lemma follows by induction on $\sigma(C)$.

THEOREM 1.

(a) *If* $\pi: C \overset{a}{\Longrightarrow} D$ *is an* (a)*-equivalence, then* π *is also a* (b)*-equivalence, and conversely.*

(b) *If there exists an* (a)*-equivalence (or a* (b)*-equivalence)* $\pi: C \Longrightarrow D$, *then* C *and* D *are formally equivalent.*

Proof. (a) The assertion is true if $\sigma^*(C) = 0$. We therefore use induction on $\sigma(C)$.

Assume that $\pi: C \Longrightarrow D$ is an (a)-equivalence. Using the notations of Definition 2, we have the (a)-equivalences $\pi_\nu': C_\nu' \to D_\nu'$, $\nu = 1, 2, \cdots, t$. Therefore, by our induction hypothesis, the π_ν' are also (b)-equivalences. So we have only to show that (using the notations of Definition 3) the pairings $\pi_\nu'^*: C_\nu' \cup \mathcal{E}' \to D_\nu' \cup E'$ are (b)-equivalences. Since $\sigma^*(C_\nu' \cup \mathcal{E}') < \sigma^*(C)$ (if $\sigma^*(C) \neq 0$), it is sufficient to show (by our induction hypothesis) that $\pi_\nu'^*$ is an (a)-equivalence. Now $\pi_\nu': C_\nu' \to D_\nu'$ is an (a)-equivalence, the curves $\mathcal{E}'$ and E' are regular, and $\pi_\nu'^*$ is an extension of π_ν', with $\pi_\nu'^*(\mathcal{E}') = E$. So, by Lemma 1, it is sufficient to show that if $\gamma_i' = \pi_\nu'(\gamma_i)$ are corresponding branches of C_ν' and D_ν', under π_ν', then $(\gamma_i', \mathcal{E}') = (\delta_i', E')$. Now, γ_i' and δ_i' are the proper transforms of the branches γ_i, δ_i of C and D respectively, and $\delta_i = \pi(\gamma_i)$. Since π is an (a)-equivalence we have $m_P(\gamma_i) = m_P(\delta_i)$, and since, by (3), we have $m_P(\gamma_i) = (\gamma_i', \mathcal{E}')$ and $m_Q(\delta_i) = (\delta_i', E')$, the equality $(\gamma_i', \mathcal{E}') = (\delta_i', E')$ is proved.

Conversely, assume that $\pi: C \to D$ is a (b)-equivalence. Then, by induction, the (b)-equivalences $\pi_\nu': C_\nu' \to D_\nu'$ are also (a)-equivalences. There thus remains only to show that if $\delta_i = \pi(\gamma_i)$ then $m_P(\gamma_i) = m_Q(\delta_i)$. Now, consider the (b)-equivalences $\pi_\nu'^*: C_\nu' \cup \mathcal{E}' \to D_\nu' \cup E'$ (which are extensions of the π_ν'). By induction, π_ν^* is also an (a)-equivalence. Hence, by Lemma 2, we have $(\gamma_i', \mathcal{E}') = (\delta_i', E')$ for any $i = 1, 2, \cdots, h$. This implies, by (3), that $m_P(\gamma_i) = m_Q(\delta_i)$, as asserted.

From now on we may replace the terms (a)-equivalence and (b)-equiv-

alence by the single term of *equivalence* (to be distinguished, for the moment, from *formal equivalence*).

(b) Assume that there exists an equivalence $\pi\colon C \underset{=}{\Longrightarrow} D$. We have then the equivalences $\pi_\nu\colon C_\nu' \underset{=}{\Longrightarrow} D_\nu'$ and $\pi_\nu'^*\colon C_\nu' \cup \mathcal{E}' \underset{=}{\Longrightarrow} D_\nu' \cup E'$. Hence, by induction on $\sigma^*(C)$, it follows that C_ν' and D_ν' are formally equivalent and that also $C_\nu' + \mathcal{E}'$ and $D_\nu' + E'$ are formally equivalent. Thus C and D are formally equivalent.

4 bis. Continuation: proof that formal equivalence $C \overset{f}{\equiv} D$ implies the existence of a pairing equivalence $\pi\colon C \underset{=}{\Longrightarrow} D$.

LEMMA 3. *Let C and D be algebroid curves having the same number of branches, let $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ denote certain branches of C and assume that these branches ε_α are regular and have distinct tangents. Similarly, let $e_1, e_2, \cdots, e_m$ be regular branches of D, with distinct tangents. Let*

$$C = \Gamma \cup \varepsilon_1 \cup \varepsilon_2 \cup \cdots \cup \varepsilon_m, \qquad D = \Delta \cup e_1 \cup e_2 \cup \cdots \cup e_m.$$

Assume that Γ and Δ are equivalent and that for each $\alpha = 1, 2, \cdots, m$ also $\Gamma \cup \varepsilon_1 \cup \varepsilon_2 \cup \cdots \cup \varepsilon_\alpha$ and $\Delta \cup e_1 \cup e_2 \cup \cdots \cup e_\alpha$ are equivalent. Then there exists a pairing equivalence $\pi\colon C \underset{=}{\Longrightarrow} D$ such that $\pi(\varepsilon_\alpha) = e_\alpha$, for $\alpha = 1, 2, \cdots, m$.

Proof. We consider two cases, according as $m = 1$ or $m > 1$.

Case I. $m = 1$. We divide this case into two subcases, according as *the tangent of ε* $(= \varepsilon_1)$ is not or is one of the tangents of C.

Case Ia. $C = \Gamma \cup \varepsilon$, $D = \Delta \cup e$, and the tangent of ε is not a tangent of Γ. Since $\Gamma \equiv \Delta$, Γ and Δ have the same number of distinct tangent lines. Similarly, $\Gamma \cup \varepsilon \equiv \Delta \cup e$ implies that $\Gamma \cup \varepsilon$ and $\Delta \cup e$ have the same number or distinct tangent lines. But since the number of distinct tangents of $\Gamma \cup \varepsilon$ is one greater than the number of distinct tangents of Γ, the same is true of $\Delta \cup e$ and Δ. *Hence the tangent of e is different from any tangent of Δ.* It follows at once from the definition of (a)-equivalence that given any equivalence $\sigma\colon \Gamma \underset{=}{\Longrightarrow} \Delta$, the extended pairing $\pi\colon \Gamma \cup \varepsilon \underset{=}{\Longrightarrow} \Delta \cup e$ defined by $\pi(\varepsilon) = e$ is also an equivalence.

Case Ib. The tangent of ε coincides with one of the tangents of Γ (and hence the tangent of e coincides with one of the tangents of Δ, by the reasoning of the Case Ia).

Let $\Gamma_1, \Gamma_2, \cdots, \Gamma_t$ and $\Delta_1, \Delta_2, \cdots, \Delta_t$ be the tangential components of Γ and Δ respectively, arranged in some arbitrary order, except that we assume

that the tangent of ε coincides with the tangent of Γ_1 and the tangent of e coincides with the tangent of Δ_1. Then the tangential components of C are $\Gamma_1 \cup \varepsilon, \Gamma_2, \cdots, \Gamma_t$, and the tangential components of D are $\Delta_1 \cup e, \Delta_2, \cdots, \Delta_t$. We shall first show that

$$(5) \qquad \Gamma_1 \equiv \Delta_1 \text{ and } \Gamma_1 \cup \varepsilon \equiv \Delta_1 \cup e.$$

Let a be the number of curves in the set $\{\Gamma_1 \cup \varepsilon, \Gamma_2, \cdots, \Gamma_t\}$ which are equivalent to Γ_1 $(a \geqq 0)$. Since $\Gamma_1 \cup \varepsilon \not\equiv \Gamma_1$, it follows that $a+1$ is the number of curves in the set $\{\Gamma_1, \Gamma_2, \cdots, \Gamma_t\}$ which are equivalent to Γ_1. Since $\Gamma \equiv \Delta$, any Γ_i is equivalent to at least one Δ_j, and any Δ_i is equivalent to at least one Γ_j. Since equivalence of algebroid curves is obviously a transitive relation it follows that $a+1$ is also the number of curves in the set $\{\Delta_1, \Delta_2, \cdots, \Delta_t\}$ which are equivalent to Γ_1. Similarly, from $\Gamma \cup \varepsilon = \Delta + e$ follows that a is the number of curves in the set $\{\Delta_1 + e, \Delta_2, \cdots, \Delta_t\}$ which are equivalent to Γ_1. Therefore, we must have necessarily $\Gamma_1 \equiv \Delta_1$.

Similarly, if we denote by b the number of curves in this set $\{\Gamma_1, \Gamma_2, \cdots, \Gamma_t\}$ which are equivalent to $\Gamma_1 \cup \varepsilon$ $(b \geqq 0)$, then $b+1$ is the number of curves in the set $\{\Gamma_1 \cup \varepsilon, \Gamma_2, \cdots, \Gamma_t\}$ which are equivalent to $\Gamma_1 \cup \varepsilon$. Since $\Gamma \equiv \Delta$ and $\Gamma \cup \varepsilon \equiv \Delta + e$, it follows, by a similar argument as above, that $\Gamma_1 \cup \varepsilon \equiv \Delta_1 \cup e$.

We have thus shown that the assumptions of the lemma are satisfied when Γ and Δ are replaced by Γ_1 and Δ_1 respectively (and thus C and D are replaced by $\Gamma_1 \cup \varepsilon$ and $\Delta_1 \cup e$ respectively). With Γ and Δ replaced by Γ_1 and Δ_1 we haves a special case of the lemma, namely the case in which $m=1$ and *each of the two curves C and D has only one tangent.* Assume that the lemma has already been proved in this special case. Then we can assert that there exist an equivalence $\pi_1 \colon \Gamma_1 \cup \varepsilon \to \Delta_1 \cup e$ such that $\pi_1(\varepsilon) = e$. We shall show now that π_1 can be extended to an equivalence $\pi \colon C \to D$. Consider the set $\{\Gamma_1 \cup \varepsilon, \Gamma_2, \cdots, \Gamma_t\}$ of tangential component of C and the set $\{\Delta_1 \cup e, \Delta_2, \cdots, \Delta_t\}$ of tangential components of D. Let M denote the set of those tangential components of C which are equivalent to $\Gamma_1 \cup \varepsilon$ (M contains at least one element, namely $\Gamma_1 \cup \varepsilon$). Similarly let N be the set of those tangeneial components of D which are equivalent to $\Delta_1 \cup e$. Since $\Gamma_1 \cup \varepsilon \equiv \Delta_1 \cup e$ and since equivalence is transitive, it follows that M and N have the same number of elements. Now, by assumption, we have $C \equiv D$, i.e., there exists some equivalence $\rho \colon C \underset{=}{\Longrightarrow} D$. Fix such an equivalence ρ. For $\nu = 1, 2, \cdots, t$, ρ induces an equivalence $\rho_\nu \colon C_\nu \underset{=}{\Longrightarrow} D_\nu$, where $C_1 = \Gamma_1 \cup \varepsilon$, $C_\nu = \Gamma_\nu$ for $\nu = 2, \cdots, t$, and $D_1, D_2, \cdots, D_t$ are the curves $\Delta_1 \cup e, \Delta_2, \cdots, \Delta_t$, in some order. We have already our equivalence $\pi_1 \colon \Gamma_1 \cup \varepsilon \underset{=}{\Longrightarrow} \Delta_1 \cup e$. We define π_ν, for $\nu = 2, \cdots, t$ as follows: a) If C_ν $(= \Gamma_\nu) \notin M$ (i.e., if

$\Gamma_\nu \not\equiv \Gamma_1 \cup \varepsilon$), then clearly also $D_\nu \notin N$ (since $\Gamma_\nu \equiv D_\nu$ and $\Gamma_1 \cup \varepsilon \equiv \Delta_1 \cup e$, whence $D_\nu \not\equiv \Delta_1 \cup e$); in this case we set $\pi_\nu = \rho_\nu \colon \Gamma_\nu \Longrightarrow D_\nu$). b) We set up an arbitrary (1, 1) correspondence between the set of curves Γ_μ in M. *other than* $\Gamma_1 \cup \varepsilon$, and the set of curves Δ_μ in N, *other than* $\Delta_1 + e$. Assuming that corresponding curves are furnished with the same index μ, we have that $\Gamma_\mu \equiv \Delta_\mu$ for all μ such that $\Gamma_1 + \varepsilon \neq \Gamma_\mu \in M$ and $\Delta_1 + e \neq \Delta_\mu \in N$ (since $\Gamma_\mu \equiv \Gamma_1 \cup \varepsilon$, $\Delta_\mu \equiv \Delta_1 \cup e$ and $\Gamma_1 \cup \varepsilon \equiv \Delta_1 \cup e$). We then fix an *arbitrary equivalence* $\pi_\mu \colon \Gamma_\mu \Longrightarrow \Delta_\mu$. This defines π_ν for all $\nu = 1, 2, \cdots, t$. Since $\Gamma_1 \cup \varepsilon, \Gamma_2, \cdots, \Gamma_t$ and $\Delta_1 \cup e, \Delta_2, \cdots, \Delta_t$ are the tangential components of C and D respectively, it follows that the ν equivalences π_ν patch up together to an equivalence $\pi \colon C \Longrightarrow D$ (use for instance the definition of (a)-equivalence).

Thus, in order to complete the proof of the lemma (in the case $m = 1$) we have only to show that (in the case Ib under consideration) there exists an equivalence $\pi_1 \colon \Gamma_1 \cup \varepsilon \Longrightarrow \Delta_1 \cup e$ such that $\pi_1(\varepsilon) = e$. Let us apply our quadratic transformations T and S (with center P and Q respectively) to $\Gamma_1 \cup \varepsilon$ and $\Delta_1 \cup e$. If $\mathcal{E}'$ and E' are the exceptional curves of T and S, we will have the total transforms

$$T\{\Gamma_1 \cup \varepsilon\} = \Gamma_1' \cup \varepsilon' \cup \mathcal{E}',$$
$$S\{\Delta_1 \cup e\} = \Delta_1' \cup e' \cup E',$$

where $\Gamma_1' = T(\Gamma_1)$, $\varepsilon' = T(\varepsilon)$; $\Delta_1' = S(\Delta_1)$, $e' = S(e)$. Since $\Gamma_1 \equiv \Delta_1$, it follows that $\Gamma_1' \equiv \Delta_1'$ (by the definition of (a)-equivalence). Since $\Gamma_1 \cup \varepsilon \equiv \Delta_1 \cup e$, it follows that $\Gamma_1' \cup \varepsilon' \equiv \Delta_1' \cup e'$ and that

$$\Gamma_1' \cup \varepsilon' \cup \mathcal{E}' \equiv \Delta_1' \cup e' \cup E'$$

(by the definition of (b)-equivalence). Furthermore, ε' and $\mathcal{E}'$ are regular curves, with distinct tangents; similarly, e' and E' are regular curves, with distinct tangents. We have here therefore the case $m = 2$ of our lemma. Since $\sigma^*(\Gamma_1 \cup \varepsilon) < \sigma^*(C)$ (unless $e^*(C) = 0$, in which case the lemma is trivially true), we may assume, by induction, that there exists an equivalence

$$\pi_1^* \colon \Gamma_1' \cup \varepsilon' \cup \mathcal{E}' \Longrightarrow \Delta_1' \cup e' \cup E'$$

such that $\pi_1^*(\varepsilon') = e'$ and $\pi_1^*(\mathcal{E}') = E'$. Then π_1^* induces an equivalence $\pi_1 \colon \Gamma_1 \cup \varepsilon \Longrightarrow \Delta_1 \cup e$ such that $\pi_1(\varepsilon) = e$.

Case II. $m > 1$. Let C_j be the tangential component of C determined by the tangent of ε_j (C_j may be ε_j itself). Similarly let D_j be the tangential component of D determined by the tangent of e_j. Then $C_1, C_2, \cdots, C_m$ are

distinct tangential components of C, and similarly for $D_1, D_2, \cdots, D_m$ and D. Since $\Gamma \cup \varepsilon_1 \cup \cdots \cup \varepsilon_{j-1} \equiv \Delta \cup e_1 \cup \cdots \cup e_{j-1}$ and

$$\Gamma \cup \varepsilon_1 \cup \cdots \cup \varepsilon_{j-1} \cup \varepsilon_j \equiv \Delta \cup e_1 \cup e_2 \cup \cdots \cup e_{j-1} \cup e_j,$$

it follows, by the case $m = 1$ that there exists an equivalence

$$\sigma_j\colon \Gamma \cup \varepsilon_1 + \cdots \cup \varepsilon_{j-1} \cup \varepsilon_j \underset{\equiv}{\longrightarrow} \Delta \cup e_1 \cup \cdots \cup e_{j-1} \cup e_j$$

such that $\sigma_j(\varepsilon_j) = e_j$. Then σ_j induces an equivalence $\pi_j\colon C_j \underset{\equiv}{\longrightarrow} D_j$. In the course of the proof of the lemma in the case I we have shown that *if* $C \equiv D$ and if $\pi_1\colon C_1 \underset{\equiv}{\longrightarrow} D_1$ is a *given* equivalence between two tangential components C_1 and D_1 of C and D respectively, then π_1 can be extended to an equivalence $\pi\colon C \underset{\equiv}{\longrightarrow} D$. In a similar fashion it can be shown that the m equivalences $\pi_j\colon C_j \underset{\equiv}{\longrightarrow} D_j$ can be extended to (i.e., are induced by) a single equivalence $\pi\colon C \underset{\equiv}{\longrightarrow} D$. Since $\pi_j(\varepsilon_j) = e_j$, the proof of the lemma is complete.

COROLLARY. *If two algebroid curves C and D are formally equivalent, they are equivalent.*

In the notations of Definition 4 we have, by hypothesis, that C_ν' and D_ν' are formally equivalent, and also that $C_\nu' \cup \mathcal{E}'$ and $D_\nu' \cup E'$ are formally equivalent $(\nu = 1, 2, \cdots, t)$. By induction on $\sigma^*(C)$ we may therefore assume that $C_\nu' \equiv D_\nu'$ and $C_\nu' \cup \mathcal{E}' \equiv D_\nu' \cup E'$. Hence, by the preceding lemma, there exists an equivalence $\pi_\nu'\colon C_\nu' \cup \mathcal{E}' \longrightarrow D_\nu' \cup E'$ such that $\pi_\nu'(\mathcal{E}') = E'$. But that means that C and D are equivalent (namely (b)-equivalent).

Note. In a Lincei Note ("La risoluzione delle singolarità delle superficie algebriche immerse," Nota I. Accademia Nazionale dei Lincei, Rendiconti, vol. XXXI, fasc. 3-4, Settembre-Ottobre (1961)) we gave the following inductive definition of equivalence of algebroid singularities (we use the notatians of Definition 2, § 3):

$C \equiv D$ if the following three conditions are satisfied:

1) $m_P(C) = m_Q(D)$.

2) C and D have the same number t of distinct tangents.

3) For a suitable ordering of the connected components $C_1', C_2', \cdots, C_t'$ and $D_1', D_2', \cdots, D_t'$ of the proper quadratic transforms of C and D, it is true that $C_\nu' \equiv D_\nu'$ $(\nu = 1, 2, \cdots, t)$.

This definition is not "correct." In fact, even the following, more exacting, definition of equivalence would not be "correct":

$C \equiv D$ if there exist a tangentially stable pairing $\pi\colon C \to D$ such that

1′) $m_P(\gamma_i) = m_Q(\delta_i)$, if $\delta_i = \pi(\gamma_i)$.

2′) $C_\nu' \equiv D_\nu'$ for $\nu = 1, 2, \cdots, t$, where C_ν' and D_ν' are the proper transforms of tangential components C_ν and D_ν which are associated with each other under π.

Here is a counter-example.

Let C consist of two branches γ_1, γ_2, having the same tangent $(y = 0)$, where γ_1 and γ_2 are defined by the following power series:

$$\gamma_1\colon x = t^4, \qquad y = t^{10} + t^{11};$$
$$\gamma_2\colon x = t^6 + t^9, \qquad y = t^{10} + t^{13}.$$

Let also D consist of the following two branches δ_1, δ_2, with the same tangent $x = 0$:

$$\delta_1\colon x = t^{10} + t^{13}, \qquad y = t^4;$$
$$\delta_2\colon x = t^{10} + t^{11}, \qquad y = t^6 + t^7.$$

We have $m_P(\gamma_1) = m_P(\delta_1) = 4$ and $m_P(\gamma_2) = m_P(\delta_2) = 6$. Hence the pairing $\pi\colon C \to D$ defined by $\pi(\gamma_1) = \delta_1$ and $\pi(\gamma_2) = \delta_2$ satisfies condition (1′) above. Applying to C the quadratic transformation $T\colon x' = x$, $y' = y/x$, we get that $C' = T(C) = \gamma_1' \cup \gamma_2'$, where $\gamma_1' = T(\gamma_1)$ and $\gamma_2' = T(\gamma_2)$ are defined by

$$\gamma_1'\colon x' = t^4, \qquad y' = t^6 + t^7;$$
$$\gamma_2'\colon x' = t^6 + t^9, \qquad y' = t^4.$$

Similarly, applying to D the quadratic transformation $S\colon x' = x/y$, $y' = y$, we get $D' = S(D) = \delta_1' \cup \delta_2'$, where $\delta_1' = S(\delta_1)$ and $\delta_2' = S(\delta_2)$ are defined by

$$\delta_1'\colon x' = t^6 + t^9, \qquad y' = t^4;$$
$$\delta_2'\colon x' = t^4, \qquad y' = t^6 + t^7.$$

Thus $\gamma_1' = \delta_2'$ and $\gamma_2' = \delta_1'$, whence $C' \equiv D'$. However, C and D are not equivalent in any reasonable algebro-geometric sense. This is so, because already the corresponding branches γ_1, δ_1 (both having a 4-fold point at P) are not equivalent. This can be seen by observing that the 4th quadratic transform $\gamma_1^{(4)}$ of γ_1 is a regular curve, while the 4th quadratic transform $\delta_1^{(4)}$ of δ_1 has a cusp. (Similarly it can be seen that γ_2, δ_2 are not equivalent).

5. Some auxiliary results on derivations. In this section we deal with an *integrally closed* (Noetherian) *local domain* R, which we assume to

be *pseudo-geometric* (see Nagata, *Local Rings*, p. 131), and with the quotient field K of R.

We consider a finite separable algebraic extension K' of K, and we denote by R' the integral closure of R in K'. Thus R' is a finite R-module. (Later on, in this section, we shall consider the more general case of a ring R' which is a finite R-module, is integrally closed in its total ring of quotients and *is free from nilpotent elements.*)

THEOREM 2. *Let D be a derivation of K, with values in K, which is regular in R (i.e., such that $DR \subset R$), and let D' be the unique extension of D to K'. Assume the following:*

A. *If $\mathfrak{p}'$ is any minimal prime ideal of R' (i.e., $\mathfrak{p}'$ is a prime ideal of hight 1 in R'), then $\mathfrak{p}'$ is tamely ramified over R.*

B. *If $\mathfrak{p}'$ is any prime ideal of R', of hight 1, which is ramified over R and if $\mathfrak{p}' \cap R = \mathfrak{p}$, then there exists an element x in $\mathfrak{p}$ such that $x \notin \mathfrak{p}^{(2)}$ and $Dx = 0$.*

Under these assumptions we have $D'R' \subset R'$.

[*Note.* Following Abhyankar ("Tame coverings and fundamental groups of algebraic varieties, Part I," *American Journal of Mathematics*, vol. 81, no. 1 (1959), p. 53) we say that $\mathfrak{p}'$ is tamely ramified over R if the following is true:

Let $\mathfrak{p} = \mathfrak{p}' \cap R$, let $(S, \mathfrak{M})$ and $(S', \mathfrak{M}')$ be respectively the completions of the localizations $R_{\mathfrak{p}}$ and $R_{\mathfrak{p}'}'$, and let F, F' be the fields of quotients of S and S' respectively. (Note that since R is pseudogeometric, S and S' are integral domains.) Let F^* be the least Galois extensions of F which contains F' (note that since K' is separable algebraic over K, also S' is separable algebraic over S). Let $(S^*, \mathfrak{M}^*)$ be the integral closure of S in F^*. (Note that since S is complete, S^* is a local domain). Then

$$[F^* : F] \cdot [S^*/\mathfrak{M}^* : S/\mathfrak{M}]_s^{-1} \not\equiv 0 \pmod{p},$$

where p is the characteristic of S/M. Note that this condition is vacuous if $p = 0$ and that "$\mathfrak{p}'$ unramified" implies "$\mathfrak{p}'$ tamely ramified."]

Proof. Since $R' = \cap R_{\mathfrak{p}'}'$, where the intersection symbol is extended to all minimal prime ideals $\mathfrak{p}'$ of R', it is sufficient to show that $D'R'_{\mathfrak{p}'} \subset R'_{\mathfrak{p}'}$ for every minimal prime ideal $\mathfrak{p}'$ of R'. Now if $\mathfrak{p}' \cap R = \mathfrak{p}$, then $\mathfrak{p}$ is a minimal prime ideal in R; and if $\mathfrak{p}$ is any minimal prime ideal in R, then the integral closure of $R_{\mathfrak{p}}$ in K' is the intersection of all the $R'_{\mathfrak{p}'}$ such that $\mathfrak{p}' \cap R = \mathfrak{p}$. Since assumptions A and B remain valid if we replace R by $R_{\mathfrak{p}}$ ($\mathfrak{p}$-minimal in R) and R' by the integral closure of $R_{\mathfrak{p}}$ in K' we may assume that *R is a regular local ring, of dimension* 1. In this case, R' is a Dedekind domain

having only a finite number of prime ideals $\neq (0)$, all lying over the maximal ideal $\mathfrak{p}$ of R.

Let $\mathfrak{p}'$ be any of the prime ideals of R', different from (0). We wish to show that $D'R'_{\mathfrak{p}'} \subset R'_{\mathfrak{p}'}$.

The case in which $\mathfrak{p}'$ is unramified over R is trivial. Namely, we have in this case: $[F^*: F] = [S'/\mathfrak{M}': S/\mathfrak{M}]_s = [S/\mathfrak{M}': S/\mathfrak{M}] = (\text{say})\ n$, and $S'/\mathfrak{M}' = R'/\mathfrak{p}'$. If ξ is any element of R' whose $\mathfrak{M}'$-residue is a primitive element of $S'/\mathfrak{M}'$ over $S/\mathfrak{M}$, then it is immediate that $S' = S[\xi]$ (since $S'\mathfrak{p}$ is the maximal ideal of S'). The derivation D has a unique extension to S, which we shall continue to denote by D, and D has a unique extension to the field of quotients F' of S' (since ξ is separable algebraic over F), which we shall continue to denote by D'. If $f(X)$ is the minimal polynomial of ξ over F, then $f(X) \in S[X]$ and $f'(\xi) \notin \mathfrak{M}'$. Since $f'(\xi)D'\xi - f^D(\xi) = 0$, where f^D is obtained by applying D to the coefficients of f, it follows that $D'\xi \in S'$. Thus D' is regular on S'. The assertion that $D'R'_{\mathfrak{p}'} \subset R'_{\mathfrak{p}'}$ now follows from the fact that $S' \cap K' = R'_{\mathfrak{p}'}$ and that $D'K' \subset K'$ (if $a, b \in R'_{\mathfrak{p}'}$ then $S'b \cap R'_{\mathfrak{p}'} = R'_{\mathfrak{p}'} \cdot b$, and hence b divides a in S' if and only if b divides a in $R'_{\mathfrak{p}'}$).

Assume now that $\mathfrak{p}'$ is ramified over R (hence tamely ramified). Again, $S^*/\mathfrak{M}^*$ is a finite separable algebraic extension of $S/\mathfrak{M}$. Upon replacing R by the integral closure of R in the inertia field of $\mathfrak{M}^*$, and using the unramified case already settled above, we may assume that $S^*/\mathfrak{M}^* = S/\mathfrak{M}$. In this case we have $S^*\mathfrak{M} = \mathfrak{M}^{*e}$, $e \not\equiv 0 \pmod{p}$, and $[F^*: F] = e$. Let x be as in assumption B. We have then $Sx = \mathfrak{M}$ and $S^*x = \mathfrak{M}^{*e}$. Since the Galois group of F^*/F is abelian, the proof can be reduced to the case of a cyclic extension F^*/F, of degree e. The adjunction to F and F^* of a primitive e-th root of unity does not destroy the validity of assumptions A and B. We may therefore assume that F contains a primitive e-th root of unity. Since S and S^* are complete rings, it is well-known then that S^* is of the form $S[t]$, where t is an element of S^* such that $t^e = ax$, a—a unit in S. We have $et^{e-1}D't = xDa = \dfrac{t^eDa}{a}$, i.e., $D't = \dfrac{Da}{ea} \cdot t \in S^*$. Hence $D'S^* \subset S^*$, and so also $D'R'_{\mathfrak{p}'} \subset R'_{\mathfrak{p}}$. This completes the proof.

We now generalize Theorem 2 as follows:

Let R be as above, and let R' be an overring of R which is integral over R and is a finite R-module (whence R' is a semi-local ring). We assume that

(a) *R' is integrally closed in its total quotient ring;*

(b) *R' has no nilpotent elements.*

As a consequence of (a) and (b), R' is a direct sum of local, integrally closed domains:

$$R' = R'_1 \oplus R'_2 \oplus \cdots \oplus R'_h,$$

where, if $1 = e_1 + e_2 + \cdots + e_h$ is the decomposition of 1 into idempotents e_i (with $e_i e_j = 0$, if $i \neq j$), then $R'_i = R'e_i$. Let $R_i = Re_i$, and let K'_i be the field of quotients of R'_i. Then R'_i is the integral closure of R_i in K'_i and is a finite R_i-module. Furthermore, K'_i is a finite algebraic extension of the field Ke_i. We assume furthermore that

(c) *K'_i is a separable extension of Ke_i ($i = 1, 2, \cdots, h$).*

THEOREM 3. *Let D be a derivation of K, regular on R, and let D' be the unique extension of D to the total quotient ring K' ($= K'_1 \oplus K'_2 \oplus \cdots \oplus K'_h$) of R' (the uniqueness of D' follows from* (c)). *Then under the same assumption A and B of Theorem* 1, *we have $D'R' \subset R'$.*

Proof. We have the decomposition of D' into the "direct" sum of derivations e_iD' of the K'_i:

$$D' = e_1D' \oplus e_2D' \oplus \cdots \oplus e_hD'.$$

If we set $D'_i = e_iD'$ and $D_i = e_iD$, where D_i is intended as a derivation of e_iK ($D_i(e_iu) = e_iD_iu$), then D'_i is the unique extension of D_i to K'_i. Applying Theorem 2 to Re_i, $R'e_i$, we see at once that D'_i is regular on $R'e_i$, and this implies that D' is regular on R'.

The following theorem is a special case of Theorem 3:

THEOREM 4. *Let $R = k[[x_1, x_2, \cdots, x_r]]$ be a power series ring, over a field k, let K be the field of quotients of R and let.*

$$K' = K'_1 \oplus K'_2 \oplus \cdots \oplus K'_h \supset K$$

be a direct sum of fields $K'_i = K'e_i$, with K'_i a finite separable algebraic extension of Ke_i. Let R' be the integral closure of R in K'. Assume the following:

A. *Every prime ideal of R', of hight* 1, *is tamely ramified over R.*

B. *If Δ is a discriminant of a basis of K'/K consisting of elements of R', then Δ—up to a unit factor in R—is a power series which is independent of x_1.*

Under these assumptions, the derivation $\dfrac{\partial}{\partial x_1}$ of K' is regular on R'.

For the proof it is only necessary to observe that if an ideal $\mathfrak{p}'$ of R',

of hight 1, is ramified over R, then the principal ideal $\mathfrak{p}=\mathfrak{p}'\cap R$ is generated by some irreducible factor ξ of Δ. We have then $\xi\notin\mathfrak{p}^{(2)}$ and $\dfrac{\partial\xi}{\partial x_1}=0$, and thus assumption B of Theorem 3 is satisfied.

COROLLARY. *Let $R=k[[x_1,x_2,\cdots,x_r]]$ (as in Theorem 4), let $f(Z)$ be a monic separable polynomial in $R[Z]$, free from multiple factors, let K' be the total ring of quotients of the ring $R[Z]/f(Z)$, and let R' be the integral closure of R in K'. Assume that condition A of Theorem 4 is satisfied. Assume furthermore that the discriminant Δ_0 of $f(Z)$ $(\Delta_0\in R)$ is of the form ϵh, where ϵ is a unit in R and h is a power series independent of x_1. Then the derivation $\dfrac{\partial}{\partial x_1}$ is regular on R'.*

Obvious consequence of Theorem 4, since the discriminant Δ of K'/K is a divisor of Δ_0 in R.

We have in mind a certain application of Theorem 4 to complete local rings. For that application we need a lemma:

LEMMA 4. *Let $(\mathfrak{o},\mathfrak{m})$ be a complete semi-local ring of characteristic zero (with $\mathfrak{m}$ denoting the intersection of the maximal ideals of $\mathfrak{o}$) and let D be a derivation of $\mathfrak{o}$ with values in $\mathfrak{o}$. Assume that there exists an element x in $\mathfrak{m}$ of $\mathfrak{o}$ such that Dx is a unit in $\mathfrak{o}$. Then $\mathfrak{o}$ contains a ring $\mathfrak{o}_1$ of representatives of the (complete) local ring $\mathfrak{o}/\mathfrak{o}x$, having the following properties:* (a) *D is zero on $\mathfrak{o}_1$;* (b) *x is analytically independent on $\mathfrak{o}_1$;* (c) *$\mathfrak{o}$ is the power series ring $\mathfrak{o}_1[[x]]$. [It follows that x is not a zero advisor of $\mathfrak{o}$, and hence* $\dim\mathfrak{o}_1=\dim\mathfrak{o}-1$.]

Proof. Without loss of generality we may assume that $Dx=1$ (replace D by $\dfrac{1}{Dx}\cdot D$). We consider the operator

$$e^{-xD}=I-xD+\frac{x^2}{2!}D^2-\frac{x^3}{3!}D^3+\cdots,$$

where I is the identity map of $\mathfrak{o}$ and $D^{(n)}$ denotes the n-th iterate of the derivation D. It is immediate that: (1) e^{-xD} is an endomorphism of $\mathfrak{o}$; (2) if $\mathfrak{o}_1=\operatorname{Im}e^{-xD}$ then D is zero on $\mathfrak{o}$; (3) the kernel of e^{-xD} is the principal ideal $\mathfrak{o}x$ (it is obvious that kernel $\subset\mathfrak{o}x$; the opposite inclusion follows from $e^{-xD}(x)=0$). From (2) follows that: (4) the restriction of e^{-xD} to $\mathfrak{o}_1$ is the identity map. From (1), (3) and (4) it follows that $\mathfrak{o}_1$ is a ring of representatives of $\mathfrak{o}/\mathfrak{o}x$. From $\mathfrak{o}_1\cap\mathfrak{o}x=(0)$ all the remaining assertions of the lemma follow at once.

COROLLARY. *The notations and assumption being as in Theorem 4 (or*

as in its Corollary), assume furthermore that k is of characteristic zero. Then R' is a power series ring in x_1, with coefficients in a subring R'_1 of R' such that R'_1 is a ring of representatives of $R'/R'x_1$ and such that $\frac{\partial}{\partial x_1}$ is zero on R'_1.

There is one essential complement to this corollary in the case dealt with in the corollary to Theorem 4. We have namely the following:

THEOREM 5. *The notations and the assumptions being as in the Corollary to Theorem 4, assume furthermore that k is of characteristic zero. Let $A'=R[z]$ be the subring $R[Z]/(f(Z))$ of R' (here z is the f-residue of Z), let $R_1=k[[x_2,x_3,\cdots,x_r]]$, $\zeta=e^{-x_1D}(z)$ (where $D=\frac{\partial}{\partial x_1}$) and $A'_1=R_1[\zeta]$ $(=e^{-x_1D}(A'))$. Then:* (a) *the ring R'_1 $(=e^{-x_1D}R')$ is the integral closure of A'_1 (in the total quotient ring of A'_1);* (b) *if $f_0(Z)$ is the polynomial obtained from $f(Z)$ by reduction module x_1, then $f_0(Z)$ is the minimal polynomial of ζ over $k\{\{x_2,x_3,\cdots,x_r\}\}$.*

Proof. (a) Since R' is integral over A' and e^{-xD_1} is an endomorphism of R', R'_1 is integral over A'_1. From the fact that x_1 is analytically independent over R'_1 and that $R'=R'_1[[x_1]]$, follows at once that the total quotient ring K'_1 of R'_1 is a subring of the total quotient ring K' of R' and that $R'\cap K'_1=R'_1$. Hence, since R' is integrally closed in K', it follows that R'_1 is integrally closed in K'_1. To complete the proof of part (a) it remains only to show that K'_1 is also the total quotient ring of A'_1. This, however, will follow once (b) is proved. In fact, (b) implies that the total quotient ring of A'_1, as a vector space over the field $k\{\{x_2,x_3,\cdots,x_r\}\}$, has dimension n, where n is the degree of f. On the other hand, also K' has dimension n over the field $k\{\{x_1,x_2,\cdots,x_r\}\}$. This implies, in particular, that any $n+1$ element of R'_1 are linearly dependent over the ring $k[[x_1,x_2,\cdots,x_r]]$. Since x_1 is analytically independent over R'_1 (and since $k[[x_2,\cdots,x_r]]\subset R'_1$), it follows that any $n+1$ elements of R'_1 are linearly dependent over the ring $k[[x_2,x_3,\cdots,x_r]]$. Hence $\dim K'_1/k\{\{x_2,x_3,\cdots,x_r\}\}\leqq n$, and since we have just seen that the total quotient rings of A'_1 has exactly dimension n over $k\{\{x_2,x_3,\cdots,x_r\}\}$, it follows that this total quotient ring coincides with K'_1.

(b) Clearly, $f_0(\zeta)=0$, and hence the minimal polynomial $\phi(Z)$ of ζ, over $k\{\{x_2,x_3,\cdots,x_r\}\}$, divides $f_0(Z)$. On the other hand, we have $0=\phi(\zeta)=e^{-x_1D}(\phi(z))$, whence $\phi(z)\in\operatorname{Ker} e^{-x_1D}$, i.e., $\phi(z)\in R'\cdot x_1$. Now, if Δ_0 is the Z-discriminant of $f(Z)$, then $\Delta_0R'\subset A'$. Hence $\Delta_0\phi(z)=x_1g(z)$, where $g(Z)\in k[[x_1,x_2,\cdots,x_r]][Z]$, and therefore we have an identity $\Delta_0\phi(Z)-x_1g(Z)=A(Z)f(Z)$, where again, $A(Z)\in k[[x_1,x_2,\cdots,x_r]][Z]$.

Setting $x_1 = 0$ in this identity, and observing $\Delta_0(0, x_2, \cdots, x_r) \neq 0$, we see that $\phi(Z)$ is divisible by $f_0(Z)$. Hence $\phi(Z) = f_0(Z)$, as asserted.

THEOREM 6. *The assumptions and notations being as in Theorem 5, let $f(Z) = \phi_1(Z)\phi_2(Z), \cdots, \phi_h(Z)$ be the factorization of f in irreducible factors in $R[Z]$, and let $\phi_{i0}(Z)$ $(\in R_1[Z])$ be the polynomial obtained from ϕ_i by reduction modulo x_1. Then:*

(a) *$f_0(Z) = \phi_{10}(Z)\phi_{20}(Z) \cdots \phi_{h0}(Z)$ is a factorization of $f_0(Z)$ in irreducible factors in $R_1[Z]$.*

(b) *If k is an algebraically closed field (always of characteristic zero) and if $\bar{F}_1$ is an algebraic closure of $k\{\{x_1\}\}$ then the $\phi_i(Z)$ are also irreducible in $\bar{F}_1\{\{x_2, x_3, \cdots, x_r\}\}[Z]$.*

Proof. (a) The integer h is the number of direct summands of the total quotient ring K' of $R[z]$. If $1 = e_1 + e_2 + \cdots + e_h$ is the decomposition of 1 into idempotents, then from $e_i^2 = e_i$ follows $2e_iDe_i = De_i$, whence $De_i = 0$ and $e^{-x_1De_i} = e_i$, showing that the e_i belong to the total quotient ring K'_1 of A'_1 (we use here the notations of the proof of Theorem 5). Hence also K'_1 is a direct sum of h fields, and this implies that $f_0(Z)$ is a product of h distinct factors in $R_1[Z]$. This proves (a).

(b) Let $F = k\{\{x_1, x_2, \cdots, x_r\}\}$. We have $K' = F[z]$ (where z is the f-residue Z) = total quotient ring of the integral closure R' of R. Since $R' = R'_1[[x_1]]$ and $z \in R'$, it follows at once that $K' = F[\zeta]$ (note that the total quotient ring K'_1 of R'_1 is $k\{\{x_2, x_3, \cdots, x_r\}\}[\zeta]$). We now pass to any of the h direct summands $K'e_i$ of K'. We set $y = ze_i$, $\eta = \zeta e_i$ and (for simplicity) we identify $F \cdot e_i$ with F. Then $F[y] = F[Z]/(\phi_i(Z))$, $k\{\{x_2, x_3, \cdots, x_r\}\}[\eta] = k\{\{x_2, x_3, \cdots, x_r\}\}[Z]/(\phi_{i0}(Z))$, and we conclude at once that $F[y] = F[\eta]$. Let $\bar{F}$ denote the field $\bar{F}\{\{x_2, x_3, \cdots, x_r\}\}$. Our claim that the polynomial $\phi_i(Z)$ is irreducible over $\bar{F}$ is equivalent to the claim that the tensor product $F[y] \otimes_F \bar{F}$ is an integral domain. Since $F[y] = F(\eta)$, this claim is equivalent to asserting that the polynomial $\phi_{i0}(Z)$ is irreducible over $\bar{F}$. Now, the coefficients of ϕ_{i0} are in the field $k\{\{x_2, x_3, \cdots, x_r\}\}$, and ϕ_{i0} is irreducible over that field (by part (a) of the theorem). The irreducibility of ϕ_{i0} over $\bar{F}$ follows now from that fact that *$k\{\{x_2, x_3, \cdots, x_r\}\}$ is maximally algebraic in $\bar{F}$.**

* *By a simple induction, it is sufficient to prove the following: if k and K are fields, $k \subset K$, and k is maximally algebraic in K, then the power series field $k\{\{t\}\}$ is maximally algebraic in $K\{\{t\}\}$. Proof.* The natural valuation V of $K\{\{t\}\}$ (with residue field K) is the extension of the natural valuation v of $k\{\{t\}\}$ (with residue field k).

6. Specializations of algebroid curves and equivalence. If $f(x,y)$ is a power series in x, y, with coefficients in a field k, and such that $f(0,0)=0$, and if this power series is free from multiple factors and is *regular in* y, then, by the Weierstrass preparation theorem, the total quotient ring K' of $k[[x,y]]/(f)$ is a direct sum of fields which are finite algebraic extensions of the field $k\{\{x\}\}$. If n is the dimension of K', regarded as a vector space over $k\{\{x\}\}$, then $1, y, y^2, \cdots, y^{n-1}$ (or—more precisely—their f-residues) form a basis of K'; we shall denote by $\Delta^y f$ the discriminant of that basis, and we shall refer to $\Delta^y f$ as the *y-discriminant of f*. The discriminant $\Delta^y f$ is a power series in x, hence up to a unit factor—is simply a power of x. If, by the Weierstrass preparation theorem, we write $f(x,y)=\epsilon(x,y)\phi(x,y)$, where $\epsilon(x,y)$ is a unit in $k[[x,y]]$ and $\phi(x,y)$ is a polynomial in y, of degree n, with coefficients in $k[[x]]$, then $\Delta^y f$ is simply the y-discriminant of the polynomial ϕ.

Let now $f(x,y;t)$ be a power series in x, y, t, with coefficients in an algebraically closed field k, *of characteristic zero*. We assume that $f(0,0;t)$ is identically zero and the the power series f is regular in y. We denote by $\bar{F}_t$ the algebraic closure of the quotient field $k\{\{t\}\}$ of $k[[t]]$, and we regard f as a power series in x, y, with coefficients in $\bar{F}_t$. We denote by $\Delta(x,t)$ the y-discriminant of $f(\Delta(x,t)\in k[[x,t]])$ and we assume that $\Delta(x,t)$ is not identically zero. In that case f has no multiple factors (in $\bar{F}_t[[x,y]]$), and we can interpret the equation $f=0$ as defining an algebroid curve C^t over the algebraically closed field $\bar{F}_t$, with origin at $P: x=y=0$.

Since $f(x,y;t)$ has been assumed to be regular in $y, f(x,y;0)$ is not identically zero. We set $f_0(x,y)=f(x,y;0)$, so that $f_0(x,y)$ is also regular in y. We shall also assume that the discriminant $\Delta(x,t)$ is not divisible by t. In that case, the y-discriminant $\Delta(x,0)$ of f_0 is not identically zero, f_0 has no multiple factors (in $k[[x,y]]$), and the equation $f_0(x,y)=0$ defines an algebroid curve C^0 over k (and hence, *a fortiori*, also over $\bar{F}_t$), with the same origin P as C^t. We shall say that C^0 *is a specialization of* C^t *over* $t\to 0$.

Our principal object in this section is the proof of the following theorem.

THEOREM 7. *Let k be of characteristic zero (and algebraically closed), and let C^0 be a specialization of C^t over $t\to 0$. Then:*

Let Σ be a finite algebraic extension of $k\{\{t\}\}$ in $K\{\{t\}\}$, and let w be the restriction of V to Σ. Then w is an extension of v, with the same value group as v, and the residue field of w is an algebraic extension of k. Since this residue field is contained in K (= residue field of V) and since k is maximally algebraic in K, it follows that w and v have the same residue field. Since $k\{\{t\}\}$ is a complete field (with respect to the valuation v), it follows that $\Sigma = k\{\{t\}\}$.

(a) *A sufficient condition that C^0 and C^t be equivalent (in the sense of § 3) is that the y-discriminant $\Delta(x,t)$ of $f(x,y;t)$ be of the form $\epsilon(x,t)x^N$ where $\epsilon(x,t)$ is a unit in $k[[x,t]]$.*

(b) *If the line $x=0$ is not a tangent of C^0, then the above condition on $\Delta(x,t)$ is also necessary for the equivalence of C^t and C^0.*

Proof. We shall need two lemmas which we shall now state.. In order not to interrupt the proof of the theorem, we shall assume for the moment these lemmas and will prove them in the next section.

LEMMA 5. *Let $C: f(x,y)=0$ and $D: g(x,y)=0$ be two algebroid curves, over an algebraically closed ground field k of characteristic zero (with common origin $P: x=y=0$). We assume that the power series f and g are regular in both x and y, that C and D have the same number h of irreducible branches, and that the line $x=0$ is neither a tangent of C nor a tangent of D. We assume that $\Delta^x f=\Delta^x g=y^M$ (apart from unit factors in $k[[y]]$). We furthermore assume that there exists a pairing $\pi: C\rightarrow D$ between the branches $\gamma_1,\gamma_2,\cdots,\gamma_h$ of C and the branches $\delta_1,\delta_2,\cdots,\delta_h$ of D having the following properties:*

a) *If $\pi(\gamma_i)=\delta_i$ then $m_P(\gamma_i)=m_P(\delta_i)$, and the intersection numbers (l,γ_i), (l,δ_i) are equal $(i=1,2,\cdots,h)$; here l denotes the line $y=0$.*

b) *If $\phi_i(x,y)=0$ is an irreducible equation of γ_i and $\psi_i(x,y)$ is an irreducible equation of δ_i $(=\pi(\gamma_i))$, then $\Delta^x\phi_i=\Delta^x\psi_i=y^{M_i}$ (apart from unit factors in $k[[y]]$).*

Then $\Delta^y f=\Delta^y g$ (apart from unit factors in $k[[x]]$).

Note that the assumption $(l,\gamma_i)=(l,\delta_i)$ in a) can also be expressed as follows: if we assume—as we may—that the ϕ_i and ψ_i in b) are monic polynomial in x, then for each $i=1,2,\cdots,h$, the two polynomials ϕ_i and ψ_i are of the same degree.

LEMMA 6. *Let $C: f(x,y)=0$ and $D: g(x,y)=0$ be two algebroid curves, over an algebraically closed field k of characterictic zero (with common origin $P: x=y=0$). We assume that the power series f and g are regular in y and that there exists an equivalence $\pi: C \Longrightarrow D$ having the following property: if $\pi(\gamma_i)=\delta_i$ $(i=1,2,\cdots,h)$ then $(m,\gamma_i)=(m,\delta_i)$; here m denotes the line $x=0$. Then $\Delta^y f=\Delta^y g$ (apart from unit factors in $k[[x]]$).*

Note the following immediate Corollary of Lemma 6:

COROLLARY. *If two algebroid curves $C\colon f(x,y)=0$, $D\colon g(x,y)=0$ (over an algebraically closed ground field k, of characteristic zero) are equivalent and if the line $x=0$ is neither a tangent of C nor a tangent of D (whence f and g are certainly regular in y, i.e., neither f nor g is divisible by x), then $\Delta^y f=\Delta^y g$ (apart from unit factors in $k[[x]]$).*

For if $\pi\colon C \Longrightarrow D$ is any equivalence between C and D and $\pi(\gamma_i)=\delta_i$, then $(m,\gamma_i)=m_P(\gamma_i)=m_P(\delta_i)=(m,\delta_i)$ (since the line $m\colon x=0$ is neither a tangent of γ_i nor a tangent of δ_i).

We now proceed with the proof of the theorem.

(a) Our basic assumption in this part of the theorem was that the y-discriminant $\Delta(x,t)$ of $f(x,y;t)$ is of the form $\epsilon(x,t)x^N$, with $\epsilon(x,t)$ a unit in $k[[x,t]]$. This was also the basic assumption made in the corollary to Theorem 4 (with $r=2$, $x_1=t$, $x_2=x$ and $h=x^N$). Thus Theorems 5 and 6 are applicable. Theorem 6 tells us that—upon assuming that $f(x,y;t)$ is a polynomial in y—the factorization $f=\phi_1\phi_2\cdots\phi_h$ of f into irreducible factors in $k[[t,x]][y]$ is also a factorization of f in irreducible factors in $\bar{F}_t[[x]][y]$. Thus the *irreducible branches* γ_i of C are given by $\gamma_i\colon \phi_i(x,y;t)=0$. Theorem 6 also tells us that if $\phi_{i0}(x,y)=\phi_i(x,y;0)$, then the irreducible branches δ_i of C^0 are given by $\delta_i\colon \phi_{i0}(x,y)=0$. This yields then a natural *specialization pairing* $\pi\colon C^t\to C^0$ between the branches of C^t and the branches of $C^0\colon \pi(\gamma_i)=\delta_i$. Upon replacing y by $y+cx$, where c is a suitable element of k (this substitution has no effect on the discriminant $\Delta^y f$), we may assume that the line $y=0$ is not tangent to C^0 (whence f_0, and hence also f, is regular in x). We shall now show that all the assumptions of Lemma 5 are satisfied if C and D are replaced by C^t and C^0 respectively and if the roles of x and y in Lemma 5 are interchanged.

Let $\gamma\colon \phi(x,y;t)=0$ be any of the irreducible branches of C^t and let $\delta=\pi(\gamma)\colon \psi(x,y)=0$, where $\psi(x,y)=\phi(x,y;0)$. We have of course $\Delta^y\phi=\Delta^y\psi$ (apart from unit factors in $k[[x,t]]$), since $\Delta^y\phi$, as a factor of $\Delta^y f$, is still of the form $\epsilon_1(x,t)x^M$, with ϵ_1 a unit in $k[[x,t]]$, and so $\Delta^y\psi=\epsilon_1(x,0)x^M$. We pass to the local domain $k[[x,y]]/(\psi)$ of δ and we denote by η^0 the ψ residue of y and by $k(\delta)$ the field of quotients $k\{\{x\}\}[\eta^0]$ of that local ring. Similarly, we denote by η the ϕ-residue of y and by $\bar{F}^t(\gamma)$ the field of quotients $\bar{F}^t\{\{x\}\}[\eta]$ of the local ring $\bar{F}^t[[x,y]]/(\phi)$ of γ. From the proof of part (b) of Theorem 6 follows that $\bar{F}^t(\gamma)$ is obtained from $k(\delta)$ by the ground field extension $k\to\bar{F}^t$, i.e.,

$$\bar{F}^t\{\{x\}\}[\eta]=k\{\{x\}\}[\eta^0]\underset{k\{\{x\}\}}{\otimes}\bar{F}^t\{\{x\}\}.$$

If $[k(\delta):k\{\{x\}\}]=n$, then $k(\delta)=k\{\{x^{1/n}\}\}$ and $\bar{F}^t(\gamma)=\bar{F}^t\{\{x^{1/n}\}\}$. Let us denote by v_0 and v the natural valuations of the complete field $k(\delta)$ and $\bar{F}^t(\gamma)$ respectively. Setting $\xi=x^{1/n}$, we have $v_0(\xi)=v(\xi)=1$, and v is the extension of v_0.

We next show that

$$v(\eta)=v_0(\eta^0). \tag{6}$$

We know from Theorem 5 that η is given by a power series of this form

$$\eta=\eta^0+u_1t+u_2t^2+\cdots, \tag{7}$$

where the u_i are elements of $k(\delta)$ which are *integral* over $k[[x]]$, i.e., the u_i are elements of $k[[\xi]]$. Taking conjugates $\eta_1,\eta_2,\cdots,\eta_n$ of η (over $\bar{F}^t\{\{x\}\}$; this amounts to taking conjugates of $\eta^0,u_1,u_2,\cdots$ over $k\{\{x\}\}$), we observe that, for $\alpha\neq\beta$, $\eta_\alpha-\eta_\beta$ divides $\Delta^\nu f$ in $\bar{F}^t[[\xi]]$. So $\eta_\alpha-\eta_\beta$ is—apart from a unit in $\bar{F}^t[[\xi]]$—a power of ξ. Since $\eta_\alpha{}^0-\eta_\beta{}^0$ is a non-unit in $k[[\xi]]$, it follows at once that for each $i\geqq 1$, $u_{i\alpha}-u_{i\beta}$ must be divisible by $\eta_\alpha{}^0-\eta_\beta{}^0$ in $k[[\xi]]$ (here $u_{i1},u_{i2},\cdots,u_{in}$ are the conjugates of u_i over $k\{\{x\}\}$). This implies that $v_0(u_i)\geqq\min\{n,v_0(\eta^0)\}$. (Recall that we have assumed that $f(0,0;t)$ is identically zero; this implies that all the u_i are non-units in $k[[\xi]]$. Let $u_i=a_{i0}\xi^{\nu_i}+a_{i,1}\xi^{\nu_i+1}+\cdots$, $\nu_i>0$, $a_{i0}\neq 0$ $(a_{ij}\in k)$. If $\nu_i<n$, then there exists a conjugate u_{i2} such that $v_0(u_{i1}-u_{i2})=\nu_i$. So, in this case, $\nu_i\geqq v_0(\eta_1{}^0-\eta_2{}^0)\geqq v_0(\eta^0)$.). Since the line $y=0$ is not tangent to C, we have $v_0(\eta^0)\leqq v_0(x)=n$. Hence $v_0(u_i)\geqq v_0(\eta^0)$, and so $v(\eta)=v_0(\eta^0)$, as asserted.

On the other hand, we have $v_0(x)=v(x)=n$. Since

$$m_P(\gamma)=\min\{v(x),v(\eta)\} \text{ and } m_P(\delta)=\min\{v_0(x),v_0(\eta^0)\},$$

it follows from (6) that

$$m_P(\gamma)=m_P(\delta). \tag{8}$$

This holds for any two branches γ, δ such that $\delta=\pi(\gamma)$. This is part of condition a) of Lemma 5. Note also that equality (6) implies that the line $y=0$ is not tangent to any branch γ of C^t (since that line was assumed not to be a tangent of C^0). This is one of the conditions imposed in Lemma 5 (with x and y interchanged).

Note also that since $v(x)=v_0(x)=n$, we have $(l,\gamma)=(l,\delta)$, where now l is the line $x=0$. Thus both parts of condition (a) of Lemma 5 are satisfied.

It has already been pointed out above that $\Delta^\nu\phi=\Delta^\nu\psi$ (apart from unit factors in $k[[x,t]]$). Thus also condition b) of Lemma 5 is satisfied.

We can therefore conclude that $\Delta^x f = \Delta^x g$ (apart from unit factors in $k[[y,t]]$).

Since $y=0$ is not tangent to either C^t or C^0, the above conclusion tells us that in order to prove that C^t and C^0 are equivalent, *we may add to our original assumption $\Delta^y f = \Delta^y g$ the assumption that the line $x=0$ is not a tangent of C^0* (and therefore also not a tangent of C^t, in view of the fact that we have $v(x) = v_0(x)$ for any two corresponding branches γ and $\delta = \pi(\gamma)$).

The rest of the proof is fairly simple and will consist in showing, by induction on the exponent M of x in the discriminant $\Delta^y f$, that the *natural specialization pairing $\pi\colon C^t \to C^0$ introduced above is an equivalence* (if $M=0$, then both C^t and C^0 are regular curves).

First of all we show that π *is tangentially stable.* If we set $\Delta^y\phi_i = \Delta^y\phi_{i0} = x^{M_i}$ (apart from unit factors), then the expression of the discriminant as the square of the product of the differences of the roots of the polynomial and—on the other hand—the definition of intersection multiplicity given in § 4, show that $M = \sum_{i=1}^{h} M_i + \sum_{i<j} (\gamma_i, \gamma_j)^2$, where $\Delta^y f = x^M$. Similarly, $M = \sum_{i=1}^{h} M_i + \sum_{i<j} (\delta_i, \delta_j)^2$. It is also clear, since δ_i is the specialization of γ_i over $t\to 0$, that $(\gamma_i,\gamma_j) \leqq (\delta_i,\delta_j)$. Hence it follows that $(\gamma_i,\gamma_j) = (\delta_i,\delta_j)$, for all $i,j = 1,2,\cdots,h, i\neq j$. This shows that π is tangentially stable, since γ_i and γ_j (respectively δ_i and δ_j) have the same tangent if and only if $(\gamma_i,\gamma_j) > m_P(\gamma_i)m_P(\gamma_j)$ (respectively, if and only if $(\delta_i,\delta_j) > m_P(\delta_i)m_P(\delta_j)$), and we have just shown that $m_P(\gamma_i) = m_P(\delta_i)$ for all $i = 1,2,\cdots,h$.

We now apply a quadratic transformation T, with center P. Since the line $x=0$ is not tangent of either C^t or C^0, we may take

$$x' = x, \qquad y' = y/x$$

as the equations of T. Let $C_\nu'^t$, $C_\nu'^0$ be the connected components of $T(C^t)$ and $T(C^0)$ respectively. Let $(0,\alpha_\nu)$ and $(0,\alpha_{\nu 0})$ be the origins of $C_\nu'^t$ and $C_\nu'^0$ respectively. The $\alpha_{\nu 0}$ are, of course, elements of the algebraically closed field k. As to the α_ν, they are—*a priori*-elements of F^t. However—and this is an important point—the α_ν *are actually elements of the power series ring* $k[[t]]$. In fact, let $\gamma_i\colon \phi_i(x,y;t) = 0$ be any of the irreducible branches of the tangential component C_ν^t of C^t, let $s_i = m_P(\gamma_i)$ and let $\phi_{i,s_i}(x,y;t)$ be the leading form of $\phi_i(x,y;t)$. .This form has coefficients in $k[[t]]$, and the coefficient of y^{s_i} is a unit ϵ_i in $k[[t]]$. Since $\phi_{i,s_i} = \epsilon_i(y-\alpha_\nu x)^{s_i}$, it follows that $\alpha_\nu \in k[[t]]$, as asserted.

It follows that if we set $y_\nu' = y' - \alpha_\nu$, then the equation of $C_\nu'^t$ will be of the form

$$f_\nu'(x, y_\nu'; t) = 0,$$

where $f_\nu'(x, y_\nu'; t)$ is an element of $k[[x,t]][y_\nu']$, regular in y_ν'. The equation of $C_\nu'^0$ will be

$$f_{\nu 0}'(x, y_{\nu 0}') = 0. \qquad y_{\nu 0}' = y' - \alpha_\nu(0),$$

where $f_{\nu 0}'(x, Y) = f_\nu'(x, Y; 0)$, and $C_\nu'^0$ is a specialization of $C_\nu'^t$ over $t \to 0$. The product of the discriminants $\Delta^{\nu'_\nu} f_\nu'$ is a divisor of the discriminant $\Delta^\nu f$, since, if we set

$$f(x, xy'; t) = x^s f'(x, y'; t),$$

where $s = m_P(C^t)$, then $\Delta^\nu f = x^{s(s-1)} \Delta^{\nu'} f'$. Thus, the discriminant $\Delta^{\nu'_\nu} f_\nu'$ is a power of x (apart from a unit in $k[[x,t]]$. The pairing π_ν': $C_\nu'^t \to C_\nu'^0$ between the branches of $C_\nu'^t$ and the branches of $C_\nu'^0$, induced by our original specialization pairing π, is obviously still a pairing by specialization $t \to 0$. It follows, by our induction hypothesis, that each π_ν' is an equivalence. This proves that π is an equivalence.

(b) This part of Theorem 7 is an immediate consequence of the corollary to Lemma 6. For, let $\Delta^\nu f = A(x,t) x^N$, with $A(0,t) \neq 0$. Then $\Delta^\nu f_0 = A(x,0) x^N$. Since the line $x = 0$ is not tangent to C^0, it is certainly not tangent to C^t (the assumption that $C^t \equiv C^0$ implies that $m_P(C^t) = m_P(C^0)$, whence the leading form of f_0 is the specialization of the leading form of f, over $t \to 0$). It follows therefore, by the corollary to Lemma 6, that x^N is the highest power of x which divided $\Delta^\nu f$. Hence $A(0,0) \neq 0$, i.e., $A(x,t)$ is a unit in $k[[x,t]]$.

This completes the proof of Theorem 7.

7. Proofs of the Lemmas 5 and 6.

Proof of Lemma 5. We consider the equalities used already in the proof of Theorem 7:

$$M = \sum_{i=1}^{h} M_i + \sum_{i<j} (\gamma_i, \gamma_j)^2,$$

$$M = \sum M_i + \sum_{i<j} (\delta_i, \delta_j)^2.$$

They imply that

$$\sum_{i<j} (\gamma_i, \gamma_j)^2 = \sum_{i<j} (\delta_i, \delta_j)^2. \tag{9}$$

Let now (apart from unit factors in $k[[x]]$):

$$\Delta^\nu \phi_i = x^{N_i}, \qquad \Delta^\nu f = x^N;$$

$$\Delta^\nu \psi_i = x^{N_i'}, \qquad \Delta^\nu g = x^{N'}.$$

From

$$N = \sum_{i=1}^{h} N_i + \sum_{i<j} (\gamma_i, \gamma_j)^2, \tag{10}$$

$$N' = \sum_{i=1}^{h} N_i' + \sum_{i<j} (\delta_i, \delta_j)^2 \tag{10'}$$

and from (9), it follows that in order to prove that $N = N'$ it will be sufficient to prove that $N_i = N_i'$, $i = 1, 2, \cdots, h$. Now, for fixed i, let v and v' denote the natural valuations of the fields of quotients of the local rings of γ_i and δ_i respectively, and let dx, dy and $d'x$ and $d'y$ denote the differentials of x and y on γ_i and δ_i respectively. We have

$$v(dx) = v'(d'x) \ (= m_P(\gamma_i) - 1 = m_P(\delta_i) - 1), \tag{11}$$

since the line $x = 0$ is not tangent to γ_i, nor to δ_i. We also have

$$v(dy) = v'(d'y) \ (= (l, \gamma_i) - 1 = (l, \delta_i) - 1). \tag{12}$$

Finally,

$$\begin{cases} M_i = v\left(\dfrac{\partial \phi_i}{\partial x}\right) = v'\left(\dfrac{\partial \psi_i}{\partial x}\right), \\ N_i = v\left(\dfrac{\partial \phi_i}{\partial y}\right), \ N_i' = v'\left(\dfrac{\partial \psi_i}{\partial y}\right). \end{cases} \tag{13}$$

The equality $N_i = N_i'$ follows now from (11), (12), (13) and from the relations

$$\frac{\partial \phi_i}{\partial x}\, dx + \frac{\partial \phi_i}{\partial y}\, dy = 0 \qquad (\text{on } \gamma_i),$$

$$\frac{\partial \psi_i}{\partial x}\, d'x + \frac{\partial \psi_i}{\partial y}\, d'y = 0 \qquad (\text{on } \delta_i).$$

Proof of Lemma 6. This time (9) is valid because π is an equivalence (see § 4, Lemma 2). Hence, by (10) and (10′), it is sufficient to show that $N_i = N_i'$ ($i = 1, 2, \cdots, h$). Hence, we may assume that C and D are irreducible curves. We shall now proceed by induction on the numerical character $\sigma(C)$ ($= \sigma(D)$; see § 3).

First case. The line $m: x = 0$ is not tangent to C (and therefore also not tangent to D, since we have assumed that $(m, C) = (m, D)$). Without loss of generality, we may assume that $y = 0$ is tangent to both C and D.

Then the quadratic transformation T, with cented $P: x = y = 0$, gives irreducible proper transforms

$$C': f'(x, y') = 0$$
$$D': g'(x, y') = 0,$$

having origins at $x = y' = 0$ (here $y' = \frac{y}{x}$). We have $C' \equiv D'$, and if m' denotes the line $x = 0$ in the (x, y')-plane (m' is the exceptional curve of T), then we know (see (3), § 4) that $(m', C') = (m', D') = m_P(C)$ $(= m_P(D))$. Hence, by our induction hypothesis, we have $\Delta^{y'} f' = \Delta^{y'} g'$. Since

$$\Delta^{y} f = x^{s(s-1)} \Delta^{y'} f',$$
$$\Delta^{y} g = x^{s(s-1)} \Delta^{y'} g'.$$

where $s = m_P(C)$ $(= m_P(D))$, the proof in this case is complete.

Second Case: the line $x = 0$ is tangent to C (and hence also to D).

We may assume that the line $y = 0$ is not tangent to C, nor tangent to D. Then by the first case we have $\Delta^{x} f = \Delta^{x} g$ (apart from unit factors in $k[[y]]$). Using the relations

$$\frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial g}\, dy = 0 \qquad \text{on } C$$

$$\frac{\partial g}{\partial x}\, dx + \frac{\partial g}{\partial y}\, dy = 0 \qquad \text{on } D,$$

and relations analogous to (11), (12) and (13), with γ_i, δ_i replaced by C and D, and with x and y interchanged, we conclude that $\Delta^{y} f = \Delta^{y} g$.

HARVARD UNIVERSITY.

ON CLASS 2 EXTENSIONS OF ALGEBRAIC NUMBER FIELDS.[1,2]

By J. Smith.

0. Introduction. In this paper the class of a p-group G is defined in terms of a certain central series. Necessary and sufficient conditions are established for a Galois extension L/K with a class 1 Galois group to be "imbeddable" in a larger extension M/K with Galois group isomorphic to a given extension of a cyclic group of order p by the given group. An extension N/K is introduced which in a sense represents the furthest one can go along these lines without increasing the number of generators, and the Galois group $G(N/K)$ is determined. It is then shown that for a suitable sequence of fields L_n and corresponding N_n, the N_n eventually include any class 2 extension of K. By keeping track of how the $G(N_n/K)$ "fit together" it is possible for certain pairs (K,p) to determine the Galois group of the compositum of all class 2 extensions of K.

1. Basic concepts. If we are given a field K of characteristic 0, a finite extension L of K, finite groups G, $\bar{G}$, a homomorphism π of G onto $\bar{G}$ and an isomorphism $\bar{\phi}$ of $\bar{G}$ onto $G(L/K)$, the Galois group of L over K, the problem of constructing a finite extension M of L, normal over K, and an isomorphism ϕ of G onto $G(M/K)$ in such a way that $r(M/L)\circ\phi=\bar{\phi}\circ\pi$, where $r(M/L): G(M/K)\to G(L/K)$ is the restriction homomorphism, is called the *imbedding problem* $P(K,L,G,\bar{G},\pi,\bar{\phi})$ and the pair (M,ϕ) is called a solution of this problem. We note that if (M,ϕ) is a solution then ϕ takes $\ker\pi$ isomorphically onto $G(M/L)$.

We suppose in this paper that the groups G, $\bar{G}$ are p groups for some fixed prime p, in fact that $\bar{G}$ is abelian of exponent p, and that $H=\ker\pi$ is cyclic of order p and hence contained in the center of G. We further suppose that ζ, a primitive p-th root of 1, is in K. Thus $\bar{G}$ acts trivially on H by conjugation and acts trivially on U, the group of p-th roots of 1 in L, through the isomorphism $\bar{\phi}$. We fix an isomorphism $\epsilon: H\to U$, which is automatically a $\bar{G}$ isomorphism.

If K is an algebraic number field and $\mathfrak{p}$ is a prime of K we will denote

Received May 26, 1964.

[1] The work in this paper was partially supported by NSF Grant GP-379.

[2] The author wishes to express his thanks to Professor K. Iwasawa, who supervised the writing of the thesis on which this paper is based.

by $K_{\mathfrak{p}}$ and $L_{\mathfrak{p}}$ completions of K and L at the prime $\mathfrak{p}$. (Since L is normal over K, $L_{\mathfrak{p}}$ is independent of the choice of extension of $\mathfrak{p}$ to L). There is a monomorphism $\lambda_{\mathfrak{p}}\colon G(L_{\mathfrak{p}}/K_{\mathfrak{p}}) \to G(L/K)$, which is independent of the choice of extension, since L/K is abelian. If $\tilde{G}_{\mathfrak{p}}$ denotes $\bar{\phi}^{-1} \circ \lambda_{\mathfrak{p}}(G(L_{\mathfrak{p}}/K_{\mathfrak{p}}))$, $G_{\mathfrak{p}}$ denotes $\pi^{-1}(\tilde{G}_{\mathfrak{p}})$, $\pi_{\mathfrak{p}} = \pi \mid G_{\mathfrak{p}}$ and $\bar{\phi}_{\mathfrak{p}} = \bar{\phi} \mid \tilde{G}_{\mathfrak{p}}$, then the imbedding problem $P(K_{\mathfrak{p}}, L_{\mathfrak{p}}, G_{\mathfrak{p}}, \tilde{G}_{\mathfrak{p}}, \pi_{\mathfrak{p}}, \bar{\phi}_{\mathfrak{p}})$ is called the local problem at the prime $\mathfrak{p}$.

A *reduction of an imbedding problem* $P(K, L, G, \tilde{G}, \pi, \bar{\phi})$ is a problem of the form $P((K, L, G_1, \tilde{G}, \pi_1, \bar{\phi})$ where G_1 is a subgroup of G which is mapped onto $\tilde{G}$ by π and $\pi_1 = \pi \mid G_1$. Under our assumptions on G and $\ker \pi$ the given problem is reducible (admits a reduction with $G_1 \neq G$) if and only if the extension $H \to G \xrightarrow{\pi} \tilde{G}$ splits and hence G can be written as $H \oplus G_1$. It is known (see [3] Theorem 1) that $P(K, L, G, \tilde{G}, \pi, \bar{\phi})$ admits a solvable reduction if and only if each associated local problem does; in fact in view of the reciprocity law this is equivalent to the existence of a solvable reduction of the local problem for a set of primes omitting at most one prime.

For a given p-group G we let $G^{(1)} = G^p[G, G]$ and $G^{(2)} = G^{(1)p}[G, G^{(1)}]$ and say G is of *class* 1 or 2 if $G^{(1)}$ or $G^{(2)}$ respectively reduce to the identity. Clearly class 1 is equivalent to elementary abelian of type $(p, \cdots, p)$ and any class 1 group is also class 2. A group is *free of class* 2 on a set S of generators if it is of class 2 and any map of S into a class 2 group can be uniquely extended to a homomorphism of the group. Existence of such a group for a given S follows from the existence of free groups for a given S and uniqueness is trivial. We denote such a group by F_S. It is routine to verify that if $T \subset S$ then the natural injection induces a monomorphism $F_T \to F_S$ so that the subgroup of F_S generated by T is in fact free of class 2 on T. Similarly the kernel of the map $F_S \to F_T$ which is the identity on T and takes $S - T$ onto e is the smallest normal subgroup $F^N{}_{S-T}$ containing $S - T$. This implies that if $\rho \in F^N{}_{S-T}$ for all finite T then $\rho = e$.

If $S = \{\sigma_1, \cdots, \sigma_n\}$ then $C_n = F_S{}^{(1)}$ is elementary abelian and is generated by $\{(\sigma_i, \sigma_j),\ 1 \leqq i < j \leqq n$ and $\sigma_k{}^p,\ 1 \leqq k \leqq n\}$. In fact (see [5]) they are a basis for $F_S{}^{(1)}$ as a vector space over Z_p (integers mod p). Consequently the order of $F_S{}^{(1)}$ is $p^{n(n+1)/2}$. Since $F_S/F_S{}^{(1)}$ is elementary abelian with the images of the σ's as basis we see that F_S has order

$$p^{n(n+1)/2} \cdot p^n = p^{(n^2+3n)/2}.$$

PROPOSITION 1. *Let* $\tau_1, \cdots, \tau_m \in F_S$ *(S finite) be such that their images under the canonical map* $F_S \to F_S/F_S{}^{(1)}$ *are independent. Then the subgroup of* F_S *they generate is free of class* 2 *on* $T = \{\tau_1, \cdots, \tau_m\}$.

Proof. Extend the set of images of the τ's to a basis of $F_S/F_S^{(1)}$ and let $\tau_{m+1},\cdots,\tau_n$ be inverse images in F_S of the elements of $F_S/F_S^{(1)}$ which were adjoined. By the Burnside basis theorem for p-groups $\tau_1,\cdots,\tau_n$ generate F_S. Since the images in $F_S/F_S^{(1)}$ of the elements of S also form a basis there are also n elements in S. Let $F_{T'}$ be free of class 2 on n elements $\tau_1',\cdots,\tau_n'$ and let $\eta^*: F_{T'}\to F_S$ be determined by: $\eta^*(\tau_i')=\tau_i$, $i=1,\cdots,n$. Since the τ_i generate F_S, η^* is onto; since $F_{T'}$ and F_S are each free of class 2 on n generators they have the same order, hence η^* is an isomorphism. Therefore the subgroup of $F_{T'}$ generated by $\tau_1',\cdots,\tau_m'$ which by a previous remark is free of class 2 on these generators is isomorphic to the subgroup of F_S generated by $\tau_1,\cdots,\tau_m$.

Now let F_n be free of class 2 on $\{\sigma_1,\cdots,\sigma_n\}$. Let $C_n=F_n^{(1)}$ and let $\bar G=F_n/C_n$. $\bar G$ has basis $\{\bar\sigma_1,\cdots,\bar\sigma_n\}$ where $\bar\sigma_i$ is the image of σ_i under the natural projection η. Let H be a cyclic group of order p on which G acts trivially (the only way it can act). Let $0\to H\xrightarrow{j} G\xrightarrow{\pi}\bar G\to 0$ be an extension of H by $\bar G$. G is then of class 2, hence there is a homomorphism $\gamma^*: F_n\to G$ such that $\pi\circ\gamma^*=\eta$. This induces a homomorphism $\gamma: C_n\to H$ which depends only on the extension $0\to H\xrightarrow{j} G\xrightarrow{\pi}\bar G\to 0$ and not on the choice of γ^*. Equivalent extensions yield the same γ, hence we obtain a map $\omega: H^2(G,H)\to \mathrm{Hom}(C_n,H)$. A simple application of MacLane's Theorem (see [4] Lemma 1) shows that ω is an isomorphism. We denote the associated map of $H^2(\bar G,U)$ onto $\mathrm{Hom}(C_n,U)=X(C_n)$ (the character group) by ω'.

Given H, $\bar G$ and C_n as above and an extension L of K and an isomorphism $\phi: \bar G\to G(L/K)$ we may associate with each $\chi\in X(C_n)$ an extension G of H by $\bar G$ which in turn yields an imbedding problem $P(K,L,G,\bar G,\pi,\bar\phi)$. We shall call a character for which the associated imbedding problem has a solvable reduction a *solvable character*. In our case this means either the problem has a solution or the extension of H by $\bar G$ splits.

2. The local problem for $\mathfrak{p}\nmid p$. If K is a $\mathfrak{p}$-adic number field with $\mathfrak{p}\nmid p$ then it is known that $[K^*: K^{*p}]=p^2$. Thus if L is a class one extension $[L:K]$ is p or p^2. The imbedding problems if $[L:K]$ is p are easy to handle; the only non-trivial one occurs if $G\approx Z_{p^2}$. This is solable if and only if L/K is unramified or K contains the p^2 roots of 1, $(p^2\mid N(\mathfrak{p})-1)$. We now consider the cases where $[L:K]=p^2$.

Case 1. $[L:K]=p^2$, $p^2\mid N(\mathfrak{p})-1$.

Let K contain ζ_k such that $\zeta_k^{p^{k-1}}=\zeta$ but no element ζ_{k+1} such that

$\zeta_{k+1}{}^p = \zeta_k$. Let π_0 be a prime element of K. Then $L = K(\zeta_{k+1}, \pi_1)$ where $\zeta_{k+1}{}^p = \zeta_k$ and $\pi_1{}^p = \pi_0$. Take $\bar{\sigma}$ and $\bar{\tau}$ in $\bar{G}$ such that $\bar{\phi}(\bar{\tau})(\zeta_{k+1}) = \zeta\zeta_{k+1}$, $\bar{\phi}(\bar{\tau})(\pi_1) = \pi_1$, $\bar{\phi}(\bar{\sigma})(\zeta_{k+1}) = \zeta_{k+1}$, $\bar{\phi}(\bar{\sigma})(\pi_1) = \zeta\pi_1$. Since $[L^* : L^{*p}]$ is also p^2 we have that there are $p+1$ cyclic extensions of order p of L, namely those gotten by adjoining a p-th root of $\zeta_{k+1}, \zeta_{k+1}\pi_1, \cdots, \zeta_{k+1}\pi_1{}^{p-1}$ and π_1. They are each normal over K and thus each provides a solution for the imbedding problem corresponding to some extension of H by $\bar{G}$. No two extensions provide solutions to the same imbedding problem, for if $\bar{\rho}$ is an element of $\bar{G}$ such that $\bar{\phi}(\bar{\rho})$ leaves a given one of $\zeta_{k+1}, \zeta_{k+1}\pi_1, \cdots, \zeta_{k+1}\pi_1{}^{p-1}$ or π_1 fixed, then the extension M gotten by adjoining the p-th root of that element of L is the only one for which any extension of $\bar{\phi}(\bar{\rho})$ to M is of order p. All of these extensions have abelian Galois groups over K.

In $H^2(\bar{G}, H)$ the cohomology classes corresponding to extensions of H by $\bar{G}$ giving imbedding problems with solvable reduction are known [2] to form a subgroup. (In fact, it is the kernel of the composite homomorphism $H^2(\bar{G}, H) \to H^2(\bar{G}, U) \to H^2(\bar{G}, L)$ where the first map is induced by ϵ, and the second by the natural injection $U \to L$). We have seen that there are at least $p+1$ elements in this subgroup and they all correspond to abelian extensions of H by $\bar{G}$. The subgroup of $H^2(\bar{G}, H)$ corresponding to the abelian extensions has order p^2, hence is exactly the set of cohomology classes corresponding to the solvable imbedding problems. If F_2 is free of class 2 on σ and τ and is mapped onto G by a homomorphism taking σ onto $\bar{\sigma}$ and τ onto $\bar{\tau}$ then looking at $X(C_2)$ we may characterize the solvable characters as those which vanish on (σ, τ), (the commutator $\sigma\tau\sigma^{-1}\tau^{-1}$).

Case 2. $[L : K] = p^2$, $p^2 \nmid N(\mathfrak{p}) - 1$.

As before the only cyclic extensions of L of degree p are gotten by adjoining p-th roots of $\zeta_{k+1}, \zeta_{k+1}\pi_1, \cdots, \zeta_{k+1}\pi_1{}^{p-1}$ and π_1. (Here, however, $k = 1$). If $x, y \in G(M/k)$ are extensions to such an M of $\bar{\phi}(\bar{\sigma})$ and $\bar{\phi}(\bar{\tau})$ respectively then computation shows that $x^p(x, y) = e$. Hence if σ', τ' are any liftings of $\bar{\sigma}$, $\bar{\tau}$ to an extension of H by $\bar{G}$ for which the imbedding problem has a solvable reduction then $\sigma'^p(\sigma', \tau') = e$. Hence if F_2, σ, τ are as before, any solvable character annihilates $\sigma^p(\sigma, \tau)$. But the solvable characters cannot be a cyclic subgroup of order p for there is at least one non-trivial solvable character which vanishes on (σ, τ) since $M = L(\zeta_{k+2})$ is an abelian extension but $G(M/K)$ is not a split extension of $G(M/L)$ by $G(L/K)$ while on the other hand $M = L(\pi_2)$, $\pi_2{}^p = \pi_1$, provides a solution to at least one imbedding problem with non-abelian G. Hence the order of the group of solvable characters is at least p^2, hence this group consists of exactly those characters which vanish on $\sigma^p(\sigma, \tau)$.

3. The global problem.

THEOREM 1. *Let K, L be algebraic number fields, such that K has only one prime dividing p and $\zeta \in K$. Let H, G, $\bar{G}$, π, $\bar{\phi}$, be as above. If the order of $\bar{G}$ is p^n form $C_n \to F_n \to \bar{G}$ as in § 1. Let χ be the character of C_n corresponding to hhe given extension of H by $\bar{G}$. For each finite prime $\mathfrak{p}$ which ramifies from K to L let $\bar{\sigma}_{\mathfrak{p}} \in \bar{G}$ be such that $\bar{\phi}(\bar{\sigma}_{\mathfrak{p}})$ leaves the spitting field of $\mathfrak{p}$ fixed, in fact leaves the maximal subfield of L on which $\mathfrak{p}$ is unramified fixed, but acts non-trivially on L, and let $\bar{\tau}_{\mathfrak{p}} \in \bar{G}$ be such that $\bar{\phi}(\bar{\tau}_{\mathfrak{p}})$ leaves fixed some subfield of L which is of degree p over the splitting field of $\mathfrak{p}$ and on which $\mathfrak{p}$ ramifies and acts on an extension of the splitting field on which $\mathfrak{p}$ remains prime (if there is one) in such a way that*

$$\lambda_{\mathfrak{p}}^{-1} \circ \bar{\phi}(\bar{\tau}_{\mathfrak{p}})(\zeta_{k+1}) = \zeta\zeta_{k+1}.$$

(ζ_{k+1} as in § 2). Let $\bar{\tau}_{\mathfrak{p}} = e$ if there is no such extension. Let $\sigma_{\mathfrak{p}}$ and $\tau_{\mathfrak{p}}$ be any lifting of $\bar{\sigma}_{\mathfrak{p}}$ and $\bar{\tau}_{\mathfrak{p}}$ to F_n.

For all real primes $\mathfrak{p}$ which become complex in L let $\bar{\phi}(\bar{\sigma}_{\mathfrak{p}})$ generate the splitting group and $\sigma_{\mathfrak{p}}$ be a lifting of $\bar{\sigma}_{\mathfrak{p}}$ to F_n.

Then the problem $P(K, L, G, \bar{G}, \pi, \bar{\phi})$ has a solvable reduction if and only if

1) *For all ramified $\mathfrak{p}$ such that $p^2 \mid N(\mathfrak{p}) - 1$, $\chi(\sigma_{\mathfrak{p}}, \tau_{\mathfrak{p}}) = 1$*
2) *For all ramified $\mathfrak{p}$ such that $p^2 \nmid N(\mathfrak{p}) - 1$, $\chi(\sigma_{\mathfrak{p}}^p(\sigma_{\mathfrak{p}}, \tau_{\mathfrak{p}})) = 1$*
3) *For all real primes which become complex in L, $\chi(\sigma_{\mathfrak{p}}^p) = 1$.*

(*Note.* This last condition can only arise when $p = 2$).

Proof. We will show that conditions 1) 2) 3) are equivalent to the solvability of a reduction of each local problem.

Consider first the infinite primes. For the complex primes and for the real primes remaining real on L, $\bar{G}_{\mathfrak{p}} = \{e\}$. Hence the local problem has a solvable reduction. For real primes which become complex on L (in case $p = 2$) we have no solutions to the local problem, hence it has a solvable reduction if and only if it is reducible, i. e. if the extension splits. This is equivalent to $\chi(\sigma^p) = 1$.

For the finite primes not dividing p which do not ramify, either $L_{\mathfrak{p}} = K_{\mathfrak{p}}$, in which case the problem admits a trivial reduction as above, or $L_{\mathfrak{p}}/K_{\mathfrak{p}}$ is of degree p and unramified and the local problem is solvable. If $\mathfrak{p} \nmid p$ and $\mathfrak{p}$ ramifies but the residue class field is not extended then $\bar{\tau}_{\mathfrak{p}} = e$, and $\bar{G}$

has order p. The local problem is reducible if and only if $\chi(\sigma_{\mathfrak{p}}^p)=1$. If $\chi(\sigma_{\mathfrak{p}}^p)\neq 1$ then by previous remarks the local problem is solvable if and only if $p^2 \mid N(\mathfrak{p})-1$. This gives conditions 1) and 2) for these primes.

For $\mathfrak{p}\nmid p$ which have both ramification and residue class field extension from K to L (i.e. $\bar{G}_{\mathfrak{p}}$ has order p^2) conditions 1) and 2) are a summary of the n.a.s.c. derived for the corresponding local problems in § 2.

This leaves only the primes dividing p. Since by assumption there is only one such the local problem corresponding to this one is solvable if and only if all the others are.[3]

4. Maximal extensions. If L/K is a finite Galois extension of class 1 and $\bar{\phi}\colon \bar{G}\to G(L/K)$, an extension N of L, normal over K is called a "*maximal class* 2 *extension of* K *with class* 1 *part* L" (abbreviated $m(K,L)$) if

1) every imbedding problem $P(K,L,G,\bar{G},\pi,\phi)$ of the above form which is solvable has a solution (M,ϕ) with $M\subset N$

2) all class 1 extensions of K contained in N are contained in L.

THEOREM 2. *Let* $K,L,\phi,\bar{G},C_n\to F_n\xrightarrow{\eta}\bar{G}$ *be as in* § 3. *Then there exists a* $m(K,L)$, N, *and a homomorphism* $\psi\colon F_n\to G(N/K)$ *such that* $\bar{\phi}\circ\eta=r(N/L)\circ\psi$. *If* X^* *denotes the set of solvable characters then* kernel $\psi=X^{*\perp}$, *those elements of* C_n *which are annihilated by* X^*.

Proof. Let $\chi_1,\cdots,\chi_m$ be a basis for X^* and let $H\to G_i\xrightarrow{\pi_i}\bar{G}$ be the corresponding extensions of H by $\bar{G}$. The resulting imbedding problems are irreducible, hence solvable. Let (M_i,ϕ_i) be solutions of these problems. Let $N=M_1M_2\cdots M_m$, let $\gamma_i^*\colon F_n\to G_i$ be such that $\pi_i\circ\gamma_i^*=\eta$, and define ψ by letting $\psi(\rho)$ act on each M_i as $\phi_i\circ\gamma_i^*(\rho)$ does for all $\rho\in F_n$. Since $r(M_i/L)\circ\phi_i\circ\gamma_i^*=\bar{\phi}\circ\pi_i\circ\gamma_i^*=\bar{\phi}\circ\eta$ the action of $\psi(\rho)$ is well defined on L and since the M_i are linearly disjoint over L, $\psi(\rho)$ is well defined on N (To verify linear disjointness it is sufficient to show that if $M_i=L(\alpha_i)$, $\alpha_i^p=a_i\in L$ then no product of the α_i with exponents $\not\equiv 0(p)$ is in L, and this follows from the fact that if such a product were in L then the product of the χ_i to the corresponding powers would be trivial, contradicting the independence of the χ_i.) Clearly

$$\ker\psi=\cap\ker\phi_i\circ\gamma_i^*=\cap\ker\chi_i=\bigcap_{\chi\in X}\ker\chi=X^{*\perp}.$$

Also $\bar{\phi}\circ\eta=r(N/L)\circ\psi$.

[3] The local problem for primes dividing p can be handled using techniques of [6] but the result is rather complicated and unsuitable for the subsequent calculations.

We next show that N is a $m(K,L)$. Suppose that χ is a solvable character with extension $H \to G \xrightarrow{\pi} \bar{G}$ and $\gamma^*: F_n \to G$. Then $\ker\chi \supset X^{*\perp}$. Let $M =$ fixed field of $\psi(\ker\chi)$. Then $r(N/M)$ factors into $\phi \circ \gamma^*$, $\phi: G \to G(M/L)$, and (M,ϕ) is a solution to the imbedding problem defined by χ.

Now let L' be a class 1 extension of K contained in N. We may assume $[L': K] = p$. If $L' \not\subset L$ then $LL' = L(\alpha)$, $\alpha^p = a \in K$. Now α must be a product of the α_i's with not all exponents $\equiv 0\,(p)$. The corresponding product of the χ_i thus gives rise to a split extension of H by $\bar{G}$ which contradicts the independence of the χ_i. This shows that N is a $m(K,L)$. We now relax the restriction that H be of order p long enough to state and prove the following two corollaries.

COROLLARY 1. *Let G be a class 2 group. Let $H = G^{(1)}$, let $\bar{G} = G/H$ and let π be the natural projection. Let F_n be a free class 2 group with the same minimal number of generators as G, let $\gamma^*: F_n \to G$ be onto and let $\pi \circ \gamma^* = \eta$. Let $\gamma: C_n \to H$ be the induced map. Let $\bar{\phi}$ be an isomorphism of $\bar{G}$ onto $G(L/K)$. Then the problem $P(K,L,G,\bar{G},\pi,\bar{\phi})$ is solvable if and only if* $\ker\gamma \supset X^{*\perp}$.

Proof. Suppose $\ker\gamma \supset X^{*\perp}$. Let N be a $m(K,L)$. Let M be the fixed field of $\psi(\ker\gamma)$ and let $\phi^* = r(N/M) \circ \psi$. ϕ^* and γ have the same kernel, hence the map ϕ^* factors to $\phi \circ \gamma$ where ϕ is an isomorphism of G onto $G(M/K)$ satisfying $r(M/L) \circ \phi = \bar{\phi} \circ \eta$. Hence (M,ϕ) is a solution.

Conversely let (M,ϕ) be a solution. For any subgroup H_1 of index p in H the extension $H/H_1 \to G/H_1 \to \bar{G}$ gives rise to an imbedding problem of the previous type which is solved by the fixed field of $\phi(H_1)$ and the natural factor map. Hence $\gamma^{-1}(H_1) \supset X^{*\perp}$. Since the intersection of all such H_1 is the identity the intersection of all $\gamma^{-1}(H_1)$ is $\ker\gamma$ which therefore contains $X^{*\perp}$.

COROLLARY 2. *Let G, $\bar{G}$, K, L, $\bar{\phi}$ be as in Corollary* 1 *and let* (M,ϕ) *be a solution of the associated problem. Then M is contained in some* $m(K,L)$.

Proof. Consider cyclic extensions M_i of L such that the M_i are linearly disjoint over L, M contains all of them and no smaller field contains all of them. Let $\phi_i = r(M/M_i) \circ \phi$ and let χ_i be a character which gives the extension $H/\ker\phi_i \to G/\ker\phi_i \to \bar{G}$. Then (M_i, ϕ_i) is a solution of the associated problem. Comparison of orders shows that the χ_i are independent

in $X(C_n)$. Extending to a basis for X^* and adding M_i corresponding to the additional χ_i's we obtain a $m(K,L)$ containing M.

Let L_{n+1} be a class 1 extension of K of degree p^{n+1}, let $\bar{G}_{n+1}$ be a class 1 group with basis $\bar{\sigma}_1,\cdots,\bar{\sigma}_{n+1}$ and let $\bar{\phi}_{n+1}\colon \bar{G}_{n+1}\to G(L_{n+1}/K)$. If $\bar{G}_n$ is a class 1 group with basis $\bar{\sigma}_1',\cdots,\bar{\sigma}_n'$, $\bar{\nu}\colon \bar{G}_{n+1}\to\bar{G}_n$ is defined by $\bar{\nu}(\bar{\sigma}_i)=\bar{\sigma}_i'$, $1\leqq i\leqq n$, $\bar{\nu}(\bar{\sigma}_{n+1})=e$, and L_n is the fixed field of $\bar{\phi}_{n+1}(\bar{\sigma}_{n+1})$ there is a unique map $\bar{\phi}_n\colon \bar{G}_n\to G(L_n/K)$ making the diagram

$$\begin{tikzcd}
\bar{G}_{n+1} \arrow[r, "\bar{\phi}_{n+1}"] \arrow[d, "\bar{\nu}"'] & G(L_{n+1}/K) \arrow[d, "r(L_{n+1}/L_n)"] \\
\bar{G}_n \arrow[r, "\bar{\phi}_n"] & G(L_n/K)
\end{tikzcd}$$

commutative.

Further if F_{n+1} is free of class 2 on $\sigma_1,\cdots,\sigma_{n+1}$, F_n free of class 2 on $\sigma_1',\cdots,\sigma_n'$ and maps ν, η_{n+1}, η_n are defined by: $\eta_{n+1}(\sigma_i)=\bar{\sigma}_i$, $1\leqq i\leqq n+1$; $\eta_n(\sigma_i')=\bar{\sigma}_i'$, $1\leqq i\leqq n$; $\nu(\sigma_i)=\sigma_i'$, $1\leqq i\leqq n$, $\nu(\sigma_{n+1})=e$, the diagram

$$\begin{tikzcd}
F_{n+1} \arrow[r, "\eta_{n+1}"] \arrow[d, "\nu"'] & \bar{G}_{n+1} \arrow[d, "\bar{\nu}"] \\
F_n \arrow[r, "\eta_n"] & \bar{G}_n
\end{tikzcd}$$

is commutative.

ν takes C_{n+1} onto C_n, hence induces a one to one map $\nu^*\colon X(C_n)\to X(C_{n+1})$ defined by $(\nu^*(\chi'))z=\chi'(\nu(z))$, $\chi'\in X(C_n)$, $z\in C_{n+1}$.

Let X_n^* be the solvable characters on C_n, i.e. those defining extensions $H\to G\xrightarrow{\pi}\bar{G}_n$ such that $P(K,L_n,G,\bar{G}_n,\pi,\bar{\phi}_n)$ has a solvable reduction. Let X_{n+1}^* be the analogous group for C_{n+1}, L_{n+1}, $\bar{\phi}_{n+1}$.

PROPOSITION 2. $\nu^*(X_n^*)=X_{n+1}^*\cap\nu^*(X(C_n))$.

Proof. The kernel of $\nu\colon C_{n+1}\to C_n$ is the subgroup generated by σ_{n+1}^p, (σ_i,σ_{n+1}) $1\leqq i\leqq n$. Therefore the image of ν^* is contained in the set of $\chi\in X(C_{n+1})$ which vanish on these elements. In fact they are equal for we can find a $\chi'\in X(C_n)$ which takes any values we like on (σ_i',σ_j') $1\leqq i<j\leqq n$ and $\sigma_k'^p$, $1\leqq k\leqq n$; hence we can make $\nu^*(\chi')$ take on any values we like on (σ_i,σ_j), $1\leqq i<j\leqq n$, σ_k^p, $1\leqq k\leqq n$, the rest of a basis

of C_{n+1}. Thus $\nu^*(X(C_n))$ are those characters for which the resulting extension is the direct product of the subgroup generated by $\gamma^*(\sigma_i)$, $1 \leqq i \leqq n$ and the group of order p generated by $\gamma^*(\sigma_{n+1})$. Hence if (M', ϕ') is a solution of the problem defined by χ' and associated map $\gamma^{*\prime}$ then $(M'L_{n+1}, \phi)$ is a solution of that defined by $\nu^*(\chi')$ and γ^* where we define ϕ by: $\phi(\gamma^*(\sigma_i))$ acts like $\phi'(\gamma^{*\prime}(\sigma_i')$ on M' and like $\tilde{\phi}_{n+1}(\tilde{\sigma}_i)$ on L_{n+1} for $1 \leqq i \leqq n$ (this is well defined since the actions coincide on $L_n = M' \cap L_{n+1}$) and $\phi(\gamma^*(\sigma_{n+1}))$ leaves M' fixed and acts like $\tilde{\phi}(\tilde{\sigma}_{n+1})$ on L_{n+1}.

Conversely if (M, ϕ) is a solution to the problem defined by $\nu^*(\chi')$ then letting $M' =$ fixed field of $\phi(\gamma^*(\sigma_{n+1}))$ and

$$\phi'(\gamma^{*\prime}(\sigma_i') = r(M/M') \circ \phi(\gamma^*(\sigma_i)), \qquad 1 \leqq i \leqq n$$

gives a solution to the problem defined by χ'.

Remark. The above result is easily seen to imply that $\nu(X_{n+1}{}^{*\perp}) = X_n{}^{*\perp}$.

PROPOSITION 3. *Given L_n, L_{n+1}, and a $m(K, L_n)$, N_n, there is a N_{n+1} which is a $m(K, L_{n+1})$ and which contains N_n.*

Proof. Let $\nu^*(\chi_1), \cdots, \nu^*(\chi_k)$ be a basis for $\nu^*(X_n{}^*)$ and choose (M_i, ϕ_i) solutions for the problems $P(K, L_n, G_i, \tilde{G}_n, \pi_i, \tilde{\phi}_n)$ determined by the χ_i, such that $M_i \subset N_n$. Then $M_1 M_2 \cdots M_k = N_n$. The $M_i L_{n+1}$ provide solutions to the problems $P(K, L_{n+1}, G_i', \tilde{G}_{n+1}, \pi_i', \tilde{\phi}_{n+1})$ defined by the $\nu^*(\chi_i)$. Extending $\{\nu^*(\chi_1), \cdots, \nu^*(\chi_k)\}$ to a basis of $X_{n+1}{}^*$ and taking (M_i, ϕ_i) $k < i \leqq k' = \dim X_{n+1}{}^*$ to be solutions of the problems determined by the additional χ_i we obtain a $m(K, L_{n+1})$ which contains N_n.

PROPOSITION 4. *If (M, ϕ), (M', ϕ') are solutions of the same imbedding problem then there is a cyclic extension L' of K of degree p such that $M' \subset ML'$.*

Proof. Consider the commutative diagram:

$$\begin{tikzcd}
G(MM'/K) \arrow[r, "r(MM'/M)"] \arrow[d, "r(MM'/M')"'] & G(M/K) \arrow[d, "r(M/L)"] \\
G(M'/K) \arrow[r, "r(M'/L)"] \arrow[ur, "\phi \circ \phi'^{-1}"] & G(L/K)
\end{tikzcd}$$

Let $G_1 = \{\gamma \in G(MM'/K) \mid \phi \circ \phi'^{-1} \circ r(MM'/M')(\gamma) = r(MM'/M)(\gamma)\}$ and let L' be the fixed field of G_1. Standard diagram chasing shows that $r(MM'/M')$ is an isomorphism of G_1 onto $G(M'/K)$. G_1 is therefore of index p, hence normal. Furthermore its intersection with $\ker r(MM'/M)$ is trivial, hence $ML' = MM' \supset M'$.

COROLLARY 1. *If N is a $m(K,L)$ and if L' contains all Galois extensions of K of degree p, then any other $m(K,L)$ is contained in NL'.*

5. Computation of $G(N/K)$. Let K be an algebraic number field of finite degree having only one prime $\mathfrak{p}$ dividing p. Let a_1 be a p^k-th root of unity in K, where K does not contain a p-th root of a_1. Let $a_2, \cdots, a_m$ be $\mathfrak{p}'$-units (units except possibly at $\mathfrak{p}$) which generate a free abelian subgroup of K^* and together with the roots of unity generate the $\mathfrak{p}'$-units. Let $a_{m+1}, \cdots, a_n$ be elements of K^* such that each a_j, $m < j \leqq n$, has a value which is not a p-th power in the value group for exactly one prime $\mathfrak{p}_j$ not dividing p, and is a unit at the other primes not dividing p. The a_i, $1 \leqq i \leqq n$ are independent $\bmod K^{*p}$ (i.e. no product with exponents $\neq 0\,(p)$ is in K^{*p}).

Let $L = K(\alpha_1, \cdots, \alpha_n)$, $\alpha_i^p = a_i$. Let $\bar{G}$ be class one with basis $\bar{\sigma}_1, \cdots, \bar{\sigma}_n$ and define $\bar{\phi}\colon \bar{G} \to G(L/K)$ by $\bar{\phi}(\bar{\sigma}_i)(\alpha_j) = \alpha_j$, $i \neq j$; $\bar{\phi}(\bar{\sigma}_i)(\alpha_i) = \zeta\alpha_i$. We shall determine the Galois group of a $m(K,L)$ over K.

For $1 \leqq i \leqq n$, $m < j \leqq n$ let $\left(\dfrac{a_i}{\mathfrak{p}_j}\right)$ be the p-th root of unity defined by $\left(\dfrac{a_i}{\mathfrak{p}_j}\right) \equiv a_i^{(N(\mathfrak{p}_j)-1)/p}\,(\mathfrak{p}_j)$.

Let $\left(\dfrac{a_i}{\mathfrak{p}_j}\right) = \zeta^{m_{ij}}$. $0 \leqq m_{ij} < p$. Then the splitting field of L at $\mathfrak{p}_j$ contains all α_i for which $m_{ij} = 0$, and if $m_{i_0j} \neq 0$ for some i_0 and $m_{i_0j}l_j \equiv 1\,(p)$ then for all i the splitting field also contains $\beta_i = \alpha_i\alpha_{i_0}^{-lm_{ij}}$.

In fact the splitting field is generated by these elements for if all the $\left(\dfrac{a_i}{\mathfrak{p}_j}\right)$ are 1 then $[L\mathfrak{p}_j : K\mathfrak{p}_j] = p$, which is equal to

$$[L : K(\alpha_1, \cdots, \alpha_{j-1}, \alpha_{j+1}, \cdots, \alpha_n)],$$

and if some $\left(\dfrac{a_i}{\mathfrak{p}_j}\right) \neq 1$ then there is a residue class field extension as well as a ramified extension and

$$[L\mathfrak{p}_j : K\mathfrak{p}_j] = p^2 = [L : K(\beta_1, \cdots, \beta_{j-1}, \beta_{j+1}, \cdots, \beta_n)]$$

(since $\beta_i = 1$).

Let $\bar{\tau}_j^* = \prod_{i \neq j} \bar{\sigma}_i^{m_{ij}}$. Then $\bar{\tau}_j^* \neq e$ and $\bar{\phi}(\bar{\tau}_j^*)$ leaves each of the $\beta_i, i \neq i_0, j$, as well as α_j fixed if $[L\mathfrak{p}_j : K\mathfrak{p}_j] = p^2$ and $\bar{\tau}_j^* = e$ if $[L\mathfrak{p}_j : K\mathfrak{p}_j] = p$. Hence $\bar{\tau}_j$, a suitable non-trivial power of $\bar{\tau}_j^*$, satisfies the conditions for the element of the splitting group denoted by $\bar{\tau}_{\mathfrak{p}_j}$ in Theorem 1. Clearly $\bar{\sigma}_j$ satisfies the conditions for $\bar{\sigma}_{\mathfrak{p}_j}$.

If there are primes $\mathfrak{p}_j$ such that $p^2 \nmid N(\mathfrak{p}_j) - 1$ then $a_1 = \zeta$. Let m_j' be such that $m_{1j} m_j' \equiv 1 (p)$, (i.e. $\left(\dfrac{\zeta}{\mathfrak{p}_j}\right)^{m_j} = \zeta$). Then if

$$\bar{\tau}_j = (\bar{\tau}_j^*)^{m_j'}, \bar{\phi}(\bar{\tau}_j)\,\zeta_2 = \bar{\phi}(\bar{\sigma}_1^{m_{1j} m_j'})\,\zeta_2 = \bar{\phi}(\bar{\sigma}_1)\,\zeta_2 = \zeta \cdot \zeta_2.$$

Thus $\bar{\phi}(\bar{\tau}_j)$ has the desired action on ζ_2 and τ_j will serve for $\bar{\tau}_{\mathfrak{p}_j}$ in Theorem 1.

If there is a real prime in K (possible only if $p = 2$) then $a_1 = -1$. For each real prime $\mathfrak{p}$ let $\bar{\sigma}_\mathfrak{p}$ denote $\prod \bar{\sigma}_i$ where the product is taken over all i such that a_i is negative at $\mathfrak{p}$. Then $\bar{\sigma}_\mathfrak{p} \neq e$ and it leaves the splitting field of $\mathfrak{p}$ in L fixed, hence satisfies the conditions of Theorem 1.

THEOREM 3. *Let K, L, $a_1, \cdots, a_n$ be as above. Let N be a $m(K,L)$ and let ψ be the map of F_n onto $G(N/K)$ given in Theorem 2. Let $\bar{\sigma}_j$ and $\bar{\tau}_j$ be as above and let σ_j and τ_j be any liftings to F_n. Then* $\ker \psi$ *is generated by*:

1) (σ_j, τ_j) *for all j such that* $p^2 \mid N(\mathfrak{p}_j) - 1$
2) $\sigma_j^p(\sigma_j, \tau_j)$ *for all j such that* $p^2 \nmid N(\mathfrak{p}_j) - 1$
3) $\sigma_\mathfrak{p}^p$ *for all real primes of K.*

Proof. A character is solvable if and only it it vanishes on all these elements, by Theorem 1. Hence they generate X^* which by Theorem 2 is $\ker \psi$.

Remark. Since $\bar{\tau}_j = (\bar{\tau}_j^*)^l$ with $p \nmid l$ and $(\sigma_j, \tau_j^{*l}) = (\sigma_j, \tau_j^*)^l$ where τ_j^* is a lifting of $\bar{\tau}_j^*$ we can use (σ_j, τ_j^*) for the generators in 1). If $p^2 \nmid N(\mathfrak{p}_j) - 1$ we have shown that l is a multiplicative inverse of $m_{1j} \pmod p$. Raising to the m_{1j}-th power $(p \nmid m_{1j})$ we see that we may take $\sigma_j^{pm_{1j}}(\sigma_j, \tau_j^*)$ for the generators in 2).

6. Infinite extensions. Let F_S be free of class 2 on a countable set of generators $S = \{\sigma_1, \sigma_2, \cdots\}$. Make F_S a topological group by taking $\{F_{S-T}{}^N \mid T \text{ finite}\}$ for a fundamental system of neighborhoods of e. Their intersection is e by a remark in § 2. One can now form the completion $\bar{F}_S$ (see [1]) which has for a fundamental system the closures $\bar{F}_{S-T}{}^N$.

Given K form a sequence of class 1 extensions L_n where $[L_n : K] = p^n$,

$L_n \subset L_{n+1}$ and any class 1 extension is contained in some L_n. For each n let N_n be a $m(K, L_n)$ such that $N_n \subset N_{n+1}$ (Proposition 3). It follows from the corollaries to Theorem 2 and Proposition 4 that any class 2 extension is contained in $N = \bigcup_n N_n$. Since N is the union of class 2 extensions of K it follows that it is the compositum of all such extensions.

Let $a_1, a_2, \cdots$ be a sequence of elements of K^* such that their images in K^*/K^{*p} form a basis. Let $L_n = K(\alpha_1, \cdots, \alpha_n)$, $\alpha_i^p = a_i$ and choose N_n as above. We wish to define a homomorphism $\bar{\psi}\colon \bar{F}_S \to G(N/K)$.

For each σ_n we let $\psi(\sigma_n)$ be any extension to N of the automorphism of $N_{n-1}L$ which leaves N_{n-1} fixed, leaves α_i fixed for $i \neq n$ and multiplies α_n by ζ. This is well defined as the description determines an automorphism on each of N_{n-1} and L which is the identity on their intersection L_{n-1}. Since $G(N/K)$ is of class 2, being the projective limit of finite class 2 groups, we can extend ψ to a homomorphism $\psi\colon F_S \to G(N/K)$.

This ψ is continuous. For the subgroups leaving the N_n fixed form a fundamental system of neighborhoods of the identity in $G(N/K)$ and if $T = \{\sigma_1, \cdots, \sigma_n\}$, F_{S-T} hence $F_{S-T}{}^N$ leaves N_n fixed. Therefore ψ can be uniquely extended to $\bar{\psi}\colon \bar{F}_S \to G(N/K)$. $\bar{\psi}$ is onto since for each N_n, $\bar{\psi}(\sigma_1), \cdots, \bar{\psi}(\sigma_n)$ generate $G(N_n/K)$ by Burnside's basis theorem, and $\bar{\psi}(\bar{F}_S)$ is closed in $G(N/K)$. Similarly ψ can be shown to be open, hence $\bar{\psi}$ is. We shall determine $\ker \bar{\psi}$ in certain special cases.

LEMMA. *If K is a finite algebraic number field such that $p \nmid h$, the class number of K, then one can find a sequence $a_1, a_2, \cdots$ of elements of K such that:*

1) $a_1 = \zeta_k$.
2) *$a_2, \cdots, a_m$ are a free basis for a group of $\mathfrak{p}'$-units which together with the roots of 1 generate all the $\mathfrak{p}'$-units.*
3) *$\{a_i K^{*p}\}$ are a basis for K^*/K^{*p}.*
4) *For each $j > m$ there exists one prime $\mathfrak{p}_j$ not dividing p such that the value of a_j at $\mathfrak{p}_j$ is not a p-th power in the value group and a_j is a unit at all other primes not dividing p.*

Proof. We can always choose $a_1, \cdots, a_m$ to satisfy 1 and 2. Index the finite primes not dividing p by the numbers $m+1, m+2, \cdots$. For each $\mathfrak{p}_j$, $j > m$ let $\mathfrak{p}_j^k$ be the lowest power of $\mathfrak{p}_j$ which is a principal ideal (this exists since h is finite). Let this ideal be (a_j). Since $p \nmid h$, $p \nmid k$, hence the value of a_j is not a p-th power in the value group. Clearly a_j is a unit at all other primes not dividing p.

To see that the a_iK^{*p} are independent suppose $a = \prod a_i^{k_i}$ is a p-th power. Its value at each finite prime $\mathfrak{p}_j$ not dividing p is a p-th power, hence $p \mid k_i$ for $i > m$. Therefore $a' = \prod_1^m a_i^{k_i}$ is a p-th power of some $\mathfrak{p}'$-unit $b = \prod_1^m a_i^{l_i}$. Therefore $pl_i = k_i$ for $1 < i \leqq m$ and $pl_1 \equiv k_1 (p^k)$. Hence $p \mid k_1$ and all exponents are $\equiv 0 (p)$.

To see the a_iK^{*p} span K^*/K^{*p} let $a \in K^*$. Since the value of each a_j at $\mathfrak{p}_j$, $m < j$ is not a p-th power, by multiplying by a suitable product, b, of a_j's we can get ab to have a value at each $\mathfrak{p}_j$ which is a p-th power. The ideal (ab) is then a p-th power of some ideal, which since $p \nmid h$ must be a prinicpal ideal (c). Hence abc^{-p} is a unit d which is a product of $a_1, \cdots, a_m$. Hence a can be written as a product of a_j's times a p-th power.

Now let K be a finite algebraic number field, containing ζ as always, having only one prime dividing p. Let $a_1, a_2, \cdots$ be a sequence satisfying conditions 1, 2, 3 and 4 of the Lemma. Construct the homomorphism $\bar{\psi}\colon \bar{F}_S \to G(N/K)$ as above.

THEOREM 4. *$\ker \bar{\psi}$ is the closed subgroup of $\bar{F}_S$ generated by:*

1) $\rho_j = \prod_{\substack{i=1 \\ i \neq j}}^{\infty} (\sigma_j, \sigma_i)^{m_{ij}}$ *for each j such that $p^2 \mid N(\mathfrak{p}_j) - 1$*

2) $\rho_j = \sigma_j^{pm_{jj}} \prod_{\substack{i=1 \\ i \neq j}}^{\infty} (\sigma_j, \sigma_i)^{m_{ij}}$ *for each j such that $p^2 \nmid N(\mathfrak{p}_j) - 1$*

3) $\rho_{\mathfrak{p}} = \prod \sigma_i^p$ *for each real prime $\mathfrak{p}$ which becomes complex in L, where the product is taken over all i such that a_i is negative at $\mathfrak{p}$. (Note: there exist such primes only if $p = 2$, $a_1 = -1$).*

Proof. First note that the infinite products converge since for any finite $T \subset S$ all but a finite number of factors are in $\bar{F}_{S-T}{}^N$ and that the product in independent of order.

For each j, $r(N/N_n) \circ \bar{\psi}(\rho_j)$ is the identity by Theorem 3, and the following remark. Likewise $r(N/N_n) \circ \bar{\psi}(\rho_{\mathfrak{p}})$ is the identity for the relevant infinite primes. Hence all the ρ's are contained in $\ker \bar{\psi}$.

Conversely suppose some $\rho \in \ker \bar{\psi}$. For any n let $T = \{\sigma_1, \cdots, \sigma_n\}$ and consider the exact sequence $0 \to \bar{F}_{S-T}{}^N \to \bar{F}_S \xrightarrow{\pi} F_T$ of continuous homomorphisms. (F_T has the discrete topology). Let $\psi_n\colon F_T \to G(N_n/K)$ be defined by $\psi_n \circ \pi = r(N/N_n) \circ \bar{\psi}$. Then the kernel of ψ_n is described by Theorem 3, and is in fact generated by the $\pi(\rho_j)$ and $\pi(\rho_{\mathfrak{p}})$. Hence for any T, ρ can be approximated modulo $\bar{F}_{S-T}{}^N$ by the ρ_j and $\rho_{\mathfrak{p}}$. Since this is true

for all n and the $\bar{F}_{S-T}{}^N$ form a fundamental system of neighborhoods of the identity ρ is in the closure of the subgroup generated by the given ρ's.

Examples.

A) $K=Q$, $p=2$. Let $a_1=-1, a_2=2, a_3, \cdots$ be the odd primes in some order. Letting $\left(\frac{a}{p}\right)$ be the usual Legendre symbol (i.e. values ± 1 for $p \nmid a$) we have:

$G(N/K)=\bar{F}_S/A$ where A is the closed subgroup generated by

1) $\prod(\sigma_j, \sigma_i)$ for all $p_j \equiv 1(4)$ where the product is taken over all i such that $\left(\frac{a_i}{p_j}\right)=-1$

2) $\sigma_j{}^2 \prod(\sigma_j, \sigma_i)$ for $p_j \equiv 3(4)$ where the product is as in 1.

3) $\sigma_1{}^2$.

B) $K=Q(i)$, $p=2$. Let $a_1=i, a_2=1+i, a_3, \cdots$ be the odd primes $\mathfrak{p}_j$ in some order. Let $\left(\frac{a_i}{\mathfrak{p}_j}\right)$ be the Legendre symbol.

Then $G(N/K)=\bar{F}_S/A$ where A is the closed subgroup generated by

1) $\prod(\sigma_j, \sigma_i)$ for all $j \geqq 3$ where the product is taken over all i such that $\left(\frac{a_i}{\mathfrak{p}_j}\right)=-1$.

More generally if K is gotten from Q by adjoining the p^k-th roots of 1, where p is a regular prime (i.e. $p \nmid$ class number of $Q(\zeta)$) it is known that the conditions of Theorem 4 hold.

UNIVERSITY OF MICHIGAN.

REFERENCES.

[1] N. Bourbaki, *Topologie Générale*, Chapter 3, § 4.

[2] R. Brauer, "Über die Konstruktion der Schiefkörper, die von endlichem Rang in bezug auf ein gegebenes Zentrum sind," *Journal für die Reine und Angewandte Mathematik*, vol. 168 (1932), pp. 44-64.

[3] H. Richter, "Über die Lösbarkeit einiger nicht-Abelscher Einbettungsprobleme," *Mathematische Annalen*, vol. 112 (1935).

[4] I. R. Šafarevič, "On the construction of fields with a given Galois group of order l^α," *American Mathematical Society Translations*, Ser. 2, Vol. 4, pp. 107-142.

[5] A. I. Skopin, "Factor groups of an upper central series of free groups," (Russian), *Doklady Akademii Nauk SSSR*, vol. 74, No. 3 (1950), pp. 425-428.

[6] ———, "p-extensions of a local field containing the $p^{m\text{th}}$ roots of unity," (Russian), *Izvestia Akademii Nauk SSSR*, vol. 19 (1955), pp. 445-470.

COMPACT FLAT RIEMANNIAN MANIFOLDS II: THE COHOMOLOGY OF Z_p-MANIFOLDS.

By L. S. CHARLAP[1] and A. T. VASQUEZ.[2]

A Z_p-manifold is a compact riemannian manifold X whose (homogeneous) holonomy group, $\Phi(X)$, is cyclic of prime order p. In [I] one of the authors classified these manifolds up to affine equivalence. In this paper we compute the cohomology of these manifolds.

It was shown in [I] that there is a 1-1 correspondence between affine equivalence classes of Z_p-manifolds and 4-tuples $(a, b, c; \mathfrak{a})$ where a, b, c are integers with $a, b > 0$; $c \geqq 0$; $c < b$; $c \leqq a$, and $\mathfrak{a}$ is an equivalence class of elments of the ideal class group of the p-th cyclotomic field $Q(\zeta_p)$. Our first result is that the cohomology of the manifold $X_p(a, b, c; \mathfrak{a})$ corresponding to $(a, b, c; \mathfrak{a})$ does not depend on $\mathfrak{a}$. In fact, if A_p is the local ring of Q at p, $(A_p = \{\frac{r}{s} \in Q : (s, p) = 1\})$, then we prove

THEOREM 1. *Let $X_p(a, b, c; \mathfrak{a})$ and $X_p(a, b, c; \mathfrak{a}')$ be Z_p-manifolds as defined above. Then there is a covering map $f : X_p(a, b, c; \mathfrak{a}) \to X_p(a, b, c; \mathfrak{a}')$ with $f^* : H^*(X_p(a, b, c; \mathfrak{a}'); G) \to H^*(X_p(a, b, c; \mathfrak{a}); G)$ a ring isomorphism for $G = Q$, Z_p or A_p. Furthermore $H^j(X_p(a, b, c; \mathfrak{a}'); Z)$ is isomorphic to $H^j(X_p(a, b, c; \mathfrak{a}); Z)$ for any j.*

Let $X = X_p(a, b, c; \mathfrak{a})$. By [I], $\dim X = (p-1)a + b$ and $H_1(X) \cong Z^b \oplus Z_p^{a-c}$ so the groups $H^j(X; Z)$ are completely determined by $\dim X$ and $H_1(X; Z)$.

The main result is the following:

THEOREM 2. *$H^j(X; Z) \cong T^j \oplus F^j$ where T^j is a vector space over Z_p of dimension d_j and F^j is a free abelian group of rank r_j. Furthermore d_j and r_j can be computed as follows:*

Expand the formal power series in t,

$$(1+t)^{b-c-1}\left(\frac{1-(\alpha t)^p}{1-\alpha t}\right)^{a-c}(1+\epsilon_p t^p)^c$$

Received June 25, 1964.

[1] Work supported by Air Force Office of Scientific Research Contract AF49(638)-253.

[2] Work supported by a National Science Foundation postdoctoral fellowship.

using the relations $\alpha^2 = 1$, $\epsilon_2 = \alpha$ *and* $\epsilon_p = 1$ *otherwise. Then* d_j *is the coefficient of* αt^{j-1}.

For r_j, *write the formal power series,*

$$(1+t)^{b-c}\left(\frac{1-(\alpha t)^p}{1-\alpha t}\right)^{a-c}(1+\epsilon_p t^p)^c$$

as $\sum_i f_i t^i + \sum_i g_i \alpha t^i$ *using the above relations. Then if* $n = \dim X$, $r_j =$

$$\frac{1}{p}\left[\binom{n}{j} + (p-1)(f_j - g_j)\right].$$

We also obtain some information about the multipicative structure of $H^*(X;Z)$ (cf. the remark following Theorem 3.2). For example, the product of any two elements of $H^*(X;Z)$ which are of finite order is zero.

The proof is based on the fact (cf. [I]) that X is a $K(\pi,1)$ and the groups π that occur are precisely those groups which satisfy a non-split exact sequence

$$(*) \qquad 0 \to M \to \pi \to Z_p \to 1$$

where M is finitely generated, free abelian and maximal abelian in π. As the notation in $(*)$ indicates, we think of M as an additive group and Z_p as a multiplicative group.

Now it follows from material in [I], that M contains a direct summand N with

$$({*\atop *}) \qquad 0 \to N \to \pi \to Z \to 1$$

exact and split. We apply the Hochschild-Serre spectral sequence to $({*\atop *})$, and it happens that $E_2 = E_\infty$.

The computation of the E_2 term of this spectral sequence follows from the determination of the exterior powers of "localized" Z_p-modules. We do this by determining the exterior powers of the indecomposable "localized" Z_p-modules and then applying a Grothendieck-type construction to obtain the general result.

As an example, in the last section we compute the cohomology of Z_p-manifolds of dimension p. If p is odd, the Poincaré polynomial of such a manifold is $\frac{1}{p}[(1+t)^p + (p-1)(1+t^p)]$.

We wish to thank J.-P. Serre for suggesting the use of "localized" modules.

1. Algebraic preliminaries. We use the term Z_p-module to denote a finitely generated free abelian group on which Z_p acts. I. Reiner in [4] has shown that isomorphism classes of Z_p-modules are in 1-1 correspondence with 4-tuples $t=(a,b,c;\mathfrak{a})$ where a, b, and c are integers satisfying $a,b,c\geqq 0$; $c\leqq a$, and $c\leqq b$, and $\mathfrak{a}$ is a member of the ideal class group, C, of the p-th cyclotomic field $Q(\zeta_p)$. We let $M(t)$ be a module in the isomorphism class determined by t.

It will be convenient to consider "localized" Z_p-modules. Let $A_p=\{\frac{r}{s}\in Q:(s,p)=1\}$, i.e. those rational numbers whose denominators are not divisible by the prime p. $A_p[Z_p]$ will denote the group ring of Z_p over the ring A_p. If M is a Z_p-module, then $\bar{M}$ will denote the $A_p[Z_p]$-module obtained from M by extending the ground ring $Z[Z_p]$ to $A_p[Z_p]$, i.e. $\bar{M}=M\otimes_Z A_p$.

LEMMA 1.1. *Let $M_i=M(a_i,b_i,c_i;\mathfrak{a}_i)$ be Z_p-modules for $i=1,2$. Then $\bar{M}_1\cong\bar{M}_2$ if $a_1=a_2$, $b_1=b_2$, and $c_1=c_2$.*

Proof. It follows from [4] (and a little work), that it suffices to show that if A_i $(i=1,2)$ are integral ideals in $\mathfrak{a}_i$, then $\bar{A}_1\cong\bar{A}_2$. We may further assume that one of these, say A_2 is R, the ring, $Z[\zeta_p]$, of algebraic integers in $Q(\zeta_p)$. Furthermore, by standard arguments (i.e. the Chinese Remainder Theorem), we may chose $A_1\in\mathfrak{a}_1$ so that the order of the group R/A_1 (i.e. the norm of the ideal A) is prime to p. Thus we have an exact sequence

$$0\to A_1\to R\to R/A_1\to 0$$

Tensoring with A_p gives the result.

Remark. Actually the converse is also true. See Proposition 3.5.

LEMMA 1.2. *Let M_i $(i=1,2)$ be Z_p-modules such that $\bar{M}_1\cong\bar{M}_2$. Then there exists a module monomorphism $\psi\colon M_1\to M_2$ with the order of coker ψ prime to p.*

Proof. Let $\phi\colon\bar{M}_1\to\bar{M}_2$ be an isomorphism. Since the M_i are finitely generated there exists an integer m, prime to p, such that

$$m\cdot\phi(M_1\otimes 1)\subset M_2\otimes 1\subset\bar{M}_2.$$

This defines a map $\psi\colon M_1\to M_2$. By choosing a basis it is clear that $\det\psi=m^N\det\phi$ where N is the rank of M_i. Since ϕ is an isomorphism this integer is prime to p; furthermore the order of coker ψ is $|\det\psi|$.

Let π be a group which satisfies (*) of the introduction. By [3] there is a spectral sequence (E_r, d_r) with

$$E_2^{i,j} \cong H^i(Z_p; H^j(M; Z)),$$

and with E_∞ isomorphic to the graded ring associated to a suitable filtration of $H^*(\pi; Z)$. Let $M^* = \operatorname{Hom}_Z(M, Z)$ with Z_p-action given by $(\sigma \cdot f)(m) = f(\sigma^{-1} \cdot m)$. Then $H^j(M; Z)$ is isomorphic as a Z_p-module to $\Lambda^j(M^*)$ where the exterior powers are taken over the *integers* and the Z_p-action is the diagonal action. Therefore to compute the E_2-term of the spectral sequence we must compute exterior powers of Z_p-modules. However, as we show in the next section, it suffices to compute exterior powers (over A_p) of the localized modules.

PROPOSITION 1.3. *Let π be a group satisfying (*). Then $H^j(\pi; Z)$ is a finitely generated abelian group; its torsion subgroup is a sum of groups of the form Z_{p^r} for some r.*

Proof. This is a trivial consequence of the existence of the above spectral sequence.

2. A localization theorem. We will call groups which satisfy (*) Z_p-Bieberbach groups.

THEOREM 2.1. *Let π_1 and π_2 be Z_p-Bieberbach groups. Thus we have the following exact sequences:*

$$(1) \qquad 0 \to M_i \to \pi_i \to Z_p \to 1 \qquad (i = 1, 2)$$

where the M_i are Z_p-modules. If $\tilde{M}_1 \cong \tilde{M}_2$, then there is a monomorphism $\psi: \pi_1 \to \pi_2$ such that $\psi^*: H^*(\pi_2; A_p) \to H^*(\pi_1; A_p)$ is a ring isomorphism.

The proof will be preceded by a technical lemma concerned with descriptions of Z_p-Bieberbach groups as extensions.

LEMMA 2.2. *Let π be a Z_p-Bieberbach group. Thus*

$$0 \to M \to \pi \to Z_p \to 1$$

is exact. This extension is described by a class $\alpha \in H^2(Z_p; M)$.

Then M has a Z_p-submodule N so that $M = Z \oplus N$ as a Z_p-module and Z_p acts trivially on the summand Z. Furthermore if $j: Z \to M$ is the inclusion, then $\alpha = j_*(\beta)$ for some $\beta \in H^2(Z_p; Z)$.

The proof of this lemma is essentially contained in Theorems 2.2 and 3.4 of [I].

Proof of Theorem 2.1. Using the lemma, we may write $M_i = Z \oplus N_i$, $j_i\colon Z \to M_i$, and $\alpha_i = (j_i)_*(\beta_i)$ $(i=1,2)$, where $\beta_i \in H^2(Z_p; Z)$. We will construct $\psi\colon \pi_1 \to \pi_2$ so that $\psi(M_1) \subset M_2$ and $\mathrm{Im}(\psi j_1) \subset \mathrm{Im}\, j_2$.

$$\begin{tikzcd} & M_1 \arrow[r] \arrow[d, "\psi \mid M_1"] & \pi_1 \arrow[d, "\psi"] \\ Z \arrow[ur, "j_1"] \arrow[r, "j_2"'] & M_2 \arrow[r] & \pi_2 \end{tikzcd}$$

Since $\tilde{M}_1 \cong \tilde{M}_2$, it follows from Lemma 1.1 that $\tilde{N}_1 \cong \tilde{N}_2$. Thus by Lemma 1.2, there is a module monomorphisms $f\colon N_1 \to N_2$ with the order of $\operatorname{coker} f$ prime to p. Because the sequences (1) are not split, $\alpha_i \neq 0$; therefore $\beta_i \neq 0$ in $H^2(Z_p; Z) \cong Z_p$. Thus there exists an integer k, prime to p, so that the following holds: Define $F\colon M_1 \to M_2$ by $F \mid N_1 = f$ and $Fj_1(n) = kj_2(n)$ for $n \in Z$. Then $F_*(\alpha_1) = \alpha_2 \in H^2(Z_p; M_2)$.

It now follows that there is a homomorphism $\psi\colon \pi_1 \to \pi_2$ making the following diagram commutative:

$$\begin{tikzcd} 0 \arrow[r] & M_1 \arrow[r] \arrow[d, "F"] & \pi_1 \arrow[r] \arrow[d, "\psi"] & Z_p \arrow[r] \arrow[d, "\text{identity}"] & 1 \\ 0 \arrow[r] & M_2 \arrow[r] & \pi_2 \arrow[r] & Z_p \arrow[r] & 1. \end{tikzcd}$$

It is clear that ψ is a monomorphism, and that $\operatorname{coker} F$ has order prime to p.

We wish to show that $\psi^*\colon H^*(\pi_2; A_p) \to H^*(\pi_1; A_p)$ is an isomorphism. By standard results, it suffices to show that $E_2^{i,j}(\psi)\colon E_2^{i,j}(\pi_2) \to E_2^{i,j}(\pi_1)$ is an isomorphism, where $(E_r(\pi_i), d_r)$ is the spectral sequence for π_i $(i=1,2)$ with coefficients in A_p. But $E_2^{i,j}(\pi_k)$ is isomorphic to $H^i(Z_p; H^j(M_k; A_p))$. Thus it suffices to show that $F^*\colon H^*(M_2; A_p) \to H^*(M_1; A_p)$ is an isomorphism.

Now M_i is a free abelian group, so we have $H^*(M_i; A_p) = \Lambda(H^1(M_i; A_p))$, the exterior algebra of $H^1(M_i; A_p)$. Hence it suffices to show that $F^*\colon H^1(M_2; A_p) \to H^1(M_1; A_p)$ is an isomorphism. This follows readily from the universal coefficient theorem and the fact that $\operatorname{coker} F$ has order prime to p so that $\operatorname{coker} F \otimes_Z A_p = 0$.

Q. E. D.

COROLLARY 2.3. *Let π be a Z_p-Bieberbach group. Then the group $H^i(\pi; Z)$ depends only on $\tilde{M} = M \otimes_Z A_p$.*

This follows from the theorem, Proposition 1.3, and the universal coefficient theorem.

Remark. i) We do not know whether the same is true of the *ring* $H^*(\pi; Z)$.

ii) Theorem 1 of the introduction now follows from 1.3 and 2.1. One can, in fact, prove the following stronger sounding result: Let $G_1, G_2, \cdots, G_k$ be any finite sequence of finite rings. Then the map f in Theorem 1 can be chosen to be a cohomology ring isomorphism with coefficients in any of the G_i or in Q or in A_p.

3. The application of the spectral sequence. We first reformulate Lemma 2.2. We use the notation of Lemma 2.2.

PROPOSITION 3.1. *N is a normal subgroup of π, and π is isomorphic to the semidirect product of N and an infinite cyclic group π/N. Furthermore π/N acts on N by projecting π/N onto π/M $(\cong Z_p)$ and using the Z_p-module structure of N.*

Proof. It is trivial that N is normal. To see that π/N is infinite cyclic note that

$$0 \to M/N \to \pi/N \to \pi/M \to 1$$

is exact, and that $M/N \cong Z$ and $\pi/M \cong Z_p$. Further, this extension is determined by $\beta \in H^2(Z_p; Z)$. As remarked in the proof of Theorem 2.1, $\beta \neq 0$; so $\pi/N \cong Z$. Since π/N is free, the sequence $0 \to N \to \pi \to \pi/N \to 1$ splits, so π is a semidirect product.

Remark. It follows from 3.1 that a Z_p-manifold is a riemannian fiber bundle over a flat circle with a flat torus as fiber and a cyclic subgroup of order p of the group of isometries of this torus as structural group.

THEOREM 3.2. *Let π be Z_p-Bieberbach group and M and N as in 3.1 (and 2.2), i.e. $M = N \oplus Z$. Then*

$$H^j(\pi; Z) \cong H^0(Z_p; \Lambda^j(M^*)) \oplus H^1(Z_p; \Lambda^{j-1}(N^*))$$

Proof. We use the spectral sequence for the exact sequence,

$$0 \to N \to \pi \to \pi/N \to 1,$$

of 3.1. Since π/N is infinite cyclic, the only groups in E_2 of total degree j which may be non-trivial are

$$E_2^{0,j} \cong H^0(\pi/N; H^j(N; Z)) \text{ and } E_2^{1,j-1} \cong H^1(\pi/N; H^{j-1}(N; Z)).$$

The bidegrees of the differentials $d_r = (d_r^{i,j})$ $(r = 2, 3, \cdots)$ show that E_∞ is isomorphic to E_2.

On the other hand $H^j(N; Z) \cong \Lambda^j(H^1(N; Z))$ as is well known, and $H^1(N; Z) \cong \mathrm{Hom}(H_1(N; Z), Z) \cong \mathrm{Hom}(N, Z)$.

The result will be an easy consequence of the following two lemmas and the fact that $H^i(\;;K)$ is an additive functor of the $Z[\;]$-module K.

LEMMA 3.3. *$\Lambda^j(M^*) \cong \Lambda^j(N^*) \oplus \Lambda^{j-1}(N^*)$ as Z_p-modules.*
The proof is trivial.

LEMMA 3.4. *Let K be a Z_p-module. Let $\rho\colon Z \to Z_p$ be an epimorphism. Then K is a $Z[Z]$ module, and*

$$H^1(Z;K) \cong H^0(Z_p;K) \oplus H^1(Z_p;K).$$

Proof. As usual, we write Z and Z_p multiplicatively. Let t be a generator of Z so $\sigma = \rho(t)$ generates Z_p. It is known that $H^1(Z;K) \cong K/L$ where L is the subgroup generated by elements of the form $l - tl$ where $l \in L$. Similarly if $\Delta = 1 - \sigma \in Z[Z_p]$ and $\Sigma = 1 + \sigma^2 + \cdots + \sigma^{p-1} \in Z[Z_p]$, it is again well known that

$$H^0(Z_p;K) \cong \operatorname{Ker}\Delta; H^1(Z_p;K) \cong \frac{\operatorname{Ker}\Sigma}{\operatorname{Im}\Delta}\;; H^2(Z_p;K) \cong \frac{\operatorname{Ker}\Delta}{\operatorname{Im}\Sigma}$$

where $\Delta\colon K \to K\colon k \to \Delta\cdot k$ and $\Sigma\colon K \to K\colon k \to \Sigma\cdot k$. Since $\Sigma\cdot\Delta = 0$ in $Z[Z_p]$, we have an exact sequence as follows:

$$0 \to \operatorname{Ker}\Sigma/\operatorname{Im}\Delta \to K/\operatorname{Im}\Delta \xrightarrow{\;\phi\;} \operatorname{Im}\Sigma \to 0$$

where ϕ is induced by Σ. Since the middle group is $H^1(Z;K)$ and $\operatorname{Im}\Sigma$ is free, we have $H^1(Z;K) \cong \operatorname{Im}\Sigma \oplus H^1(Z_p;K)$. It remains to show that $H^0(Z_p;K) \cong \operatorname{Im}\Sigma$ as an abelian group and this follows since $H^i(Z_p;K)$ is a finite group.

Remark. i) Theorem 3.2 could alternatively have been obtained from the spectral sequence for the exact sequence

$$0 \to M \to \pi \to Z_p \to 1.$$

However, in this case, E_2 is *not* E_∞, but the E_3 term can be computed by means of the results in [2], and $E_3 \cong E_\infty$. This was, in fact, our first proof of 3.2 and provided the motivation for [2] which, of course, is valid in more general situations.

ii) Some information about the multiplicative structure of $H^*(\pi;Z)$ can be obtained by examining the proof of the above theorem. There is in $H^*(\pi;Z)$ an ideal, I^*, of square zero. If we write $I^* = \bigoplus_{j=0}^{\infty} I^j$, then

$$I^j \cong H^0(Z_p;\Lambda^{j-1}(N^*)) + H^1(Z_p;\Lambda^{j-1}(N^*)).$$

In particular, I^* contains all elements of finite order; the quotient algebra, $P^* = H^*(\pi;Z)/I^*$, is isomorphic to the Z_p-invariant elements in the exterior algebra $\Lambda(N^*)$. If we consider the exact sequence

$$0 \to I^* \to H^*(\pi;Z) \to P^* \to 0,$$

we are led to a problem in the Hochschild theory of algebra extensions about which we have no information.

To make further progress on the computation of $H^j(\pi;Z)$, it is necessary to be able to compute $H^i(Z_p;\Lambda^j M^*)$, etc. in terms of the Z_p-module M. The first simplification is the following:

PROPOSITION 3.5. *Let M_1 and M_2 be Z_p-modules with $\tilde{M}_1 \cong \tilde{M}_2$. Then $H^i(Z_p;M_1) \cong H^i(Z_p;M_2)$ for all i.*

Proof. By Lemma 1.2 there is a Z_p-module monomorphism $\psi\colon M_1 \to M_2$ so that $K = \operatorname{coker}\psi$ has order prime to p; so

$$0 \to M_1 \to M_2 \to K \to 0$$

is exact. Consider the cohomology sequence of this. By [1], $H^i(Z_p;K) = 0$ for $i > 0$. Furthermore we know that $H^0(Z_p;K)$ is a finite group of order prime to p. It follows that $\psi_*\colon H^i(Z_p;M_1) \to H^i(Z_p;M_2)$ is an isomorphism for $i > 0$. We only have to see that $H^0(Z_p;M_1) \cong H^0(Z_p;M_2)$. We know, as above, that the groups are isomorphic to subgroups of M_1 and M_2 respectively and hence are free abelian groups. Since $\psi_*\colon H^0(Z_p;M_1) \to H^0(Z_p;M_2)$ is a monomorphism with finite cokernel, it follows that $H^0(Z_p;M_1)$ and $H^0(Z_p;M_2)$ are of the same rank; hence they are isomorphic.

4. The ring of localized Z_p-modules. We introduce some formalism (similar to the construction of the Grothendieck ring of an abelian category) to aid in the computation of exterior products of localized Z_p-modules.

Let S be the set of isomorphism classes of localized Z_p-modules. Let $F(S)$ be the free abelian group generated by the set S. For each pair $\tilde{M}_1$, $\tilde{M}_2$ of such modules, we define $\gamma(\tilde{M}_1,\tilde{M}_2) \in F(S)$ by the formula $\gamma(\tilde{M}_1,\tilde{M}_2) = (\tilde{M}_1 \oplus \tilde{M}_2) - (\tilde{M}_1) - (\tilde{M}_2)$. Let $R(S)$ be the subgroup of $F(S)$ generated by all such elements. Finally set $\mathcal{K} = F(S)/R(S)$.

If $\tilde{M}_1,\cdots,\tilde{M}_k$ are $A_p[Z_p]$-modules, then $\tilde{M}_1 \otimes_p \tilde{M}_2 \otimes_p \cdots \otimes_p \tilde{M}_k$ denotes the $A_p[Z_p]$-modules which as an A_p-module is $\tilde{M}_1 \otimes_{A_p} \tilde{M}_2 \otimes_{A_p} \cdots \otimes_{A_p} \tilde{M}_k$, and on which Z_p acts by the diagonal action, i.e. if $\tilde{m}_i \in \tilde{M}_i$ and $\sigma \in Z_p$, then $\sigma(\tilde{m}_1 \otimes_p \cdots \otimes_p \tilde{m}_k) = (\sigma\tilde{m}_1) \otimes_p \cdots \otimes_p (\sigma\tilde{m}_k)$. If $\tilde{M}$ is an $A_p[Z_p]$-module, then $\Lambda_p{}^i\tilde{M}$ denotes the $A_p[Z_p]$-module which as an A_p-module is $\Lambda_{A_p}{}^i\tilde{M}$ and which inherits its Z_p-action from the canonical surjection

$$\tilde{M} \otimes_{A_p} \cdots \otimes_{A_p} \tilde{M} \to \Lambda_{A_p}{}^i\tilde{M}.$$

By repeating the above definitions for Z_p-modules, we readily verify that $\tilde{M}_1 \otimes_p \cdots \otimes_p \tilde{M}_k \cong \widetilde{M_1 \otimes \cdots \otimes M_k}$ and $\Lambda_p{}^i \tilde{M} \cong \widetilde{\Lambda^i M}$. Thus the above operations take localized Z_p-modules into localized Z_p-modules.

In particular, "$\otimes_p$" gives rise to an associtive, commutative, bilinear pairing from $\mathcal{K} \times \mathcal{K}$ to $\mathcal{K}$, thus providing $\mathcal{K}$ with the structure of a commutative ring with identity—the role of the identity is played by the localization of the Z with trivial Z_p-action; we denote it by the symbol $\tilde{1} \in \mathcal{K}$.

Let $\mathcal{K}[[t]]$ denote the ring of formal power series with coefficients in the ring $\mathcal{K}$. Let $\mathcal{K}_0[[t]]$ be those elements of $\mathcal{K}[[t]]$ whose "constant" term is $\tilde{1} \in \mathcal{K}$. It is easily seen that $\mathcal{K}_0[[t]]$ is an abelian group with respect to the multiplication in $\mathcal{K}[[t]]$.

PROPOSITION 4.1. *There is a unique map* $\Lambda \colon \mathcal{K} \to \mathcal{K}[[t]]$ *such that*

i) $\Lambda(a+b) = \Lambda(a) \cdot \Lambda(b)$ *for* $a, b \in \mathcal{K}$, *and*

ii) *for each* $(\tilde{M}) \in S$,

$$\Lambda(\tilde{M}) = \tilde{1} + (\Lambda_p{}^1 \tilde{M})t + (\Lambda_p{}^2 \tilde{M})t^2 + \cdots.$$

Proof. This is essentially the well-known statement that $\Lambda_p{}^k(\tilde{M}_1 \oplus \tilde{M}_2) = \bigoplus_{i+j=k} (\Lambda_p{}^i \tilde{M}_1 \otimes_p \Lambda_p{}^j \tilde{M}_2)$. We only have to check that this A_p-module homomorphism is, in fact, an $A_p[Z_p]$-homomorphism, but this is easily verified.

To begin a study of the ring $\mathcal{K}$ and the mapping Λ, we introduce the rank homomorphism $r \colon \mathcal{K} \to Z$. A Z_p-module M is, in particular, a free abelian group, and one may associate to it the integer, $r(M)$, its rank as a Z-module. We wish to define a similar function on the set S. It is tempting to put $r(\tilde{M}) = r(M)$. This however presumes the following: If M_1 and M_2 are Z_p-modules such that $\tilde{M}_1 \cong \tilde{M}_2$, then $r(M_1) = r(M_2)$. This is, in fact, the case. To see this let p be the principal ideal in A_p generated by the integer p. Let N be an A_p-module. Then N/pN is an A_p/p-module. Now $A_p/p \cong Z_p$, so N/pN is a vector space over the field Z_p. Thus, $r(M)$ equals the dimension of the vector space $\tilde{M}/p\tilde{M}$, and hence $r(M)$ depends only on the isomorphism class of $\tilde{M}$ as an A_p-module. The function $r \colon S \to Z$ defines in a unique way a homomorphism (which we continue to call r) $r \colon \mathcal{K} \to Z$. Furthermore, we also have a homomorphism $r \colon \mathcal{K}_0[[t]] \to Z_0[[t]]$ where, by analogy with the above notation, $Z_0[[t]]$ denotes those formal power series with integer coefficients whose "constant" term is 1.

PROPOSITION 4.2. $r(\Lambda K) = (1+t)^{r(K)} \in Z_0[[t]]$ *for any* $K \in \mathcal{K}$.

Proof. Besides the formal extensions, this merely asserts that $r(\Lambda_p{}^i \tilde{M}) = \binom{r(\tilde{M})}{i}$ for $M \in S$ which is standard.

5. Computations in $\mathcal{K}$. We make the following notational conventions which only apply in this section: Denote $H^i(Z_p; M)$ by $H^i(M)$ for any Z_p-module M. If N is an $A_p[Z_p]$-module, it is, in a natural way, a $Z[Z_p]$-module, and it is with respect to this structure that $H^i(N)$ is defined. If G is an abelian group and k is an integer, then $k \cdot G$ is the abelian group $G \oplus G \oplus \cdots \oplus G$ (k times). If G has some further structure, kG is defined similarly. Note that in all other sections we write G^k for kG.

We consider some particular Z_p-modules. Denote by β the Z_p-module $Z[Z_p]$. As usual, put $\Sigma = 1 + \sigma + \sigma^2 + \cdots + \sigma^{p-1} \in \beta$ where σ is a fixed generator of Z_p. Note that $\tau\Sigma = \Sigma$ for each $\tau \in Z_p$. Hence Σ generates a Z_p-submodule of rank one on which Z_p acts trivially. We denote this Z_p-module by the symbol 1. We define the Z_p-module α by the exact sequence

$$0 \to 1 \to \beta \to \alpha \to 0. \tag{2}$$

We denote, as usual, by $\tilde{1}$, $\tilde{\alpha}$, and $\tilde{\beta}$ the respective localizations of these modules.

In terms of Reiner's invariants, we have

$$1 \cong M(0,0,1;R), \quad \alpha \cong M(1,0,0;R), \quad \text{and} \quad \beta \cong M(1,1,1;R)$$

where R is the ring of algebraic integers in $Q(\zeta_p)$, i.e. $R = Z[\zeta_p]$. Furthermore, it is clear from [4] that

$$M(a,b,c;R) = (b-c)\,1 \oplus (a-c)\,\alpha \oplus c\beta. \tag{3}$$

LEMMA 5.1.
$$H^0(1) \cong Z; H^{2i+1}(1) = 0; H^{2i}(1) \cong Z_p$$
$$H^0(\beta) \cong Z; H^{2i+1}(\beta) = 0; H^{2i}(\beta) = 0$$
$$H^0(\alpha) = 0; H^{2i+1}(\alpha) \cong Z_p; H^{2i}(\alpha) = 0$$
for $i > 0$.

Proof. The results concerning 1 and β are trivial while the cohomology of α follows from (2).

LEMMA 5.2. *Let M be a Z_p-module and $j\colon M \to \tilde{M} = M \otimes_Z A_p\colon m \to m \otimes 1$. Then $j_*\colon H^i(M) \to H^i(\tilde{M})$ is an isomorphism for $i > 0$.*

Proof. The proof follows easily from the fact that each element of $\tilde{M}/j(M)$ has order prime to p.

PROPOSITION 5.3. *With respect to its additive structure $\mathcal{K}$ is a free abelian group generated by $\tilde{1}$, $\tilde{\alpha}$, and $\tilde{\beta}$.*

Proof. The result follows easily from 1.1, 5.1, and 5.2. In fact, we see that the integer $r(\tilde{M})$ and the groups $H^i(\tilde{M})$ $(i = 1, 2)$ completely determine $\tilde{M} \in \mathcal{K}$.

The next theorem describes the multiplication in $\mathcal{K}$.

THEOREM 5.4. i) $\bar{\beta}^2 = p\bar{\beta}$
ii) $\bar{\alpha}\bar{\beta} = (p-1)\bar{\beta}$
iii) $\bar{\alpha}^2 = \bar{1} + (p-2)\bar{\beta}$.

Proof. It is actually as easy to prove $\beta \otimes \beta \cong p\beta$ and $\alpha \otimes \beta \cong (p-1)\beta$. In fact, the following more general statement holds. Let N be a Z_p-module of rank n. Then $N \otimes \beta \cong n\beta$. This follows immediately from the argument at the top of page 199 of [1] which asserts that the "diagonal action" of Z_p on $N \otimes \beta$ is equivalent to a "one-sided" action.

To prove iii), it suffices, in view of the preceding results, to show that $H^1(\bar{\alpha}^2) = 0$ and $H^2(\bar{\alpha}^2) \cong Z_p$, and this follows easily by tensoring the exact sequence (2) with α and then localizing.

Remark. We actually have proved the stronger statements i′) $\beta \otimes \beta \cong p\beta$ and ii′) $\alpha \otimes \beta \cong (p-1)\beta$. The analogue iii′) $\alpha \otimes \alpha \cong 1 \oplus (p-2)\beta$ is true, but since it is more difficult to prove, and we have no need for it later, we omit the proof.

PROPOSITION 5.5. i) $\Lambda_p{}^i\bar{\beta} = \dfrac{1}{p}\dbinom{p}{i}\bar{\beta}$ *for* $1 \leqq i \leqq p-1$

ii) $\Lambda_p{}^p\bar{\beta} = \begin{cases} \bar{1} & \text{if } p \text{ is odd} \\ \bar{\lambda} & \text{if } p \text{ is even.} \end{cases}$

Proof. For i), it suffices by 5.3 to show that $H^j(\Lambda_p{}^i\bar{\beta}) = 0$ for $j = 1, 2$ and $1 \leqq i \leqq p-1$. For this it suffices to show that $\Lambda_p{}^i\bar{\beta}$ is a direct summand of $k \cdot \bar{\beta}$ for some $\bar{\beta}$. But by definition there is an exact sequence of $A_p[Z_p]$-modules as follows:

$$0 \to K_i \to \otimes_p{}^i\bar{\beta} \xrightarrow{\rho_i} \Lambda_p{}^i\bar{\beta} \to 0.$$

We define a splitting l_i, i.e. an $A_p[Z_p]$-module homomorphism $l_i\colon \Lambda_p{}^i\bar{\beta} \otimes_p{}^i\bar{\beta}$ with $\rho_i l_i$ equal to the identity map. Let S_i be the symmetric group on i letters. S_i acts on $\otimes_p{}^i\bar{\beta}$ by permutation. We put

$$l_i(v_1 \wedge_p v_2 \wedge_p \cdots \wedge_p v_i) = \frac{1}{i!} \sum_{g \in S_i} \epsilon(g) g \cdot (v_1 \otimes_p \cdots \otimes_p v_i)$$

where $\epsilon(g)$ is the sign of the permutation g. We remark that since p is prime and $1 \leqq i \leqq p-1$, $1/i!$ is in A_p so the above formula makes sense— it is precisely at this point it becomes essential to have considered localized Z_p-modules. By 5.4, $\otimes_p{}^i\bar{\beta}$ is of the form $k \cdot \bar{\beta}$, and i) is proved.

ii) is trivial.

PROPOSITION 5.6. *For odd i in the range $1 \leqq i \leqq p-1$, we have*

$$\Lambda_p{}^i\tilde{\alpha} = \tilde{\alpha} + \frac{1}{p}\left[\binom{p-1}{i} - p + 1\right]\tilde{\beta}.$$

For even i in the range $1 \leqq i \leqq p-1$, we have

$$\Lambda_p{}^i\tilde{\alpha} = \tilde{1} + \frac{1}{p}\left[\binom{p-1}{i} - 1\right]\tilde{\beta}.$$

Proof. By the proof of 5.3, it suffices to compute $r(\Lambda_p{}^i\tilde{\alpha})$, $H^1(\Lambda_p{}^i\tilde{\alpha})$ and $H^2(\Lambda_p{}^i\tilde{\alpha})$. $r(\Lambda_p{}^i\tilde{\alpha}) = \binom{p-1}{i}$ and the other computations can be done using the exact sequence

$$0 \to \Lambda_p{}^{i-1}\tilde{\alpha} \to \Lambda_p{}^i\tilde{\beta} \to \Lambda_p{}^i\tilde{\alpha} \to 0$$

and previous results. Details are left to the reader.

We now reformulate 5.5 and 5.6. Let I be the ideal of $\mathcal{K}$ generated by $\tilde{\beta}$ (cf. 5.4. i) and ii)). Let $\rho\colon \mathcal{K} \to \mathcal{K}/I$ be the canonical surjection. Put $\bar{\mathcal{K}} = \mathcal{K}/I$ and $\bar{\mathcal{K}}_0[[t]]$ equal to the obvious thing. Let $\rho\colon \mathcal{K}_0[[t]] \to \bar{\mathcal{K}}_0[[t]]$ be the extension of ρ.

THEOREM 5.7. i) $\rho\Lambda(\tilde{1}) = \tilde{1} + t \in \bar{\mathcal{K}}_0[[t]]$

ii) $\rho\Lambda(\tilde{\alpha}) = \{\tilde{1} - (\tilde{\alpha}t)^p\}/\{\tilde{1} - \tilde{\alpha}t\}$

iii) $\rho\Lambda(\tilde{\beta}) = \begin{cases} \tilde{1} + t^p & \textit{if } p \textit{ is odd} \\ \tilde{1} + \tilde{\alpha}t^2 & \textit{if } p \textit{ is } 2. \end{cases}$

Proof. These are either trivial or immediate consequences of 5.5 and 5.6 recalling that $\alpha^2 = 1$ modulo I.

COROLLARY 5.8. $\rho\Lambda(\tilde{M}(a,b,c;R)) = (\tilde{1}+t)^{b-c}\left(\dfrac{\tilde{1}-(\tilde{\alpha}t)^p}{\tilde{1}-\tilde{\alpha}t}\right)^{a-c}(\tilde{1}+\epsilon_p t^p)^c$ *where $\epsilon_p = 1$ for $p \neq 2$ and $\epsilon_2 = \alpha$.*

Proof. Trivial from 4.1 and (3).

Theorem 2 of the introduction now follows from 5.8, 3.2 and the following proposition:

PROPOSITION 5.9. $H^j(\Lambda^i M^*) \cong H^j(\Lambda^i M)$ *for all i.*

Proof. By 3.5, it suffices to show $\tilde{M}^* \cong \tilde{M}$, and since "$*$" is an additive functor it suffices to show that $\tilde{1}^* \cong \tilde{1}$, $\tilde{\alpha}^* \cong \tilde{\alpha}$, and $\tilde{\beta}^* \cong \tilde{\beta}$. We leave this to the reader.

Remark. Theorem 2 settles completely the question of the cohomology

groups of Z_p-manifolds. One might be tempted to inquire about the characteristic classes of Z_p-manifolds. However, a theorem of J. A. Thorpe ([5]) shows that all characteristic classes (except possibly w_1, the first Stiefel-Whitney class) vanish.

6. An example. According to [I], the dimension of a Z_p-manifold must be at least p. In fact, there exists for each prime p, at least one such manifold, although in general, there are many. For $p=2$, there is only one; it is the Klein bottle. In any case, as we shall see, they all have the same cohomology groups.

PROPOSITION 6.1. *Let X_p be a p-dimensional Z_p-manifold. Then if p is odd*

$$H^i(X_p;Z) \cong Z^{c_i} + Z_p \qquad \textit{for } 2 \leqq i \leqq p-1 \textit{ and } i \textit{ even}$$

and

$$H^i(X_p;Z) \cong Z^{c_i}. \qquad \textit{for } 1 \leqq i \leqq p-2 \textit{ and } i \textit{ odd.}$$

where $c_i = 1/p\binom{p}{i}$.

Proof. It follows from [I], that the Z_p-module M corresponding to a p-dimensional Z_p-manifold has the form $M(1,1,0;\mathfrak{a})$ for some $\mathfrak{a} \subset Z[\zeta_p]$. The proposition now follows from Theorem 2.

Thus if p is odd, the Poincaré polynomial of a Z_p-manifold of dimension p is

$$\frac{(1+t)^p + (p-1)(1+t^p)}{p}.$$

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[I] L. S. Charlap, "Compact flat Riemannian manifolds I," *Annals of Mathematics*, vol. 81 (1965), p. 15.

[1] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton, 1956.

[2] L. S. Charlap and A. T. Vasquez, "The cohomology of group extensions," *to appear*.

[3] G. Hochschild and J.-P. Serre, "The cohomology of group extensions," *Transactions of the American Mathematical Society*, vol. 74 (1953), p. 110.

[4] I. Reiner, "Integral representations of cyclic groups of prime order," *Proceedings of the American Mathematical Society*, vol. 8 (1957), p. 142.

[5] J. A. Thorpe, "Parallelizability and flat manifolds," *Proceedings of the American Mathematical Society*, vol. 16 (1965), p. 138.

THE CONSTRUCTION OF MINIMAL DISCRETE FLOWS.

By Robert Ellis.[1]

An important problem in topological dynamics is: given a topological space X and a topological group T, find an action of T on X such that xT is dense $(x \in X)$, i. e. X is minimal under T. The purpose of this paper is to describe a general method for constructing such actions when X is a compact metric space and T is the integers. In particular it is shown that there exists a homeomorphism of the Klein bottle onto itself such that the Klein bottle is minimal under the resulting discrete flow.

The idea behind the above method is to start with an arbitrary action of T on X, vary this action in a systematic manner, and by means of a category argument prove the existence of a minimal action among the actions so obtained. In order to be more specific let me introduce some notation and definitions.

Let (X,T) and (Y,T) be transformation groups, π a continuous map of X onto Y. Then π is a *homomorphism of* (X,T) *onto* (Y,T) if $\pi(xt) = \pi(x)t$ $(x \in X, t \in T)$. If there exists a homomorphism of (X,T) onto (Y,T) then (X,T) *covers* (Y,T).

Let G be another topological group acting on X such that $xgt = xtg$ $(x \in X, t \in T, g \in G)$. Then T acts on the orbit space, X/G of X under G in a natural way. If (Y,T) is isomorphic to $(X/G,T)$, then (X,T) *G-covers* (Y,T).

Now suppose (X,T) G-covers (Y,T) and f is a continuous function of $Y \times T$ into G such that $f(y,ts) = f(y,t)f(yt,s)$ $(y \in Y; t,s \in T)$. Then $x * t = xtf(\pi x, t)$ $(x \in X, t \in T)$ determines a new action of T on X. If suitable restrictions are imposed on G and (Y,T), then there exist an f such that X becomes minimal under the action induced by f. For example: let X, Y be compact metric spaces, G a compact connected Lie group, T the integers, (Y,T) minimal and distal; then there exists a minimal action of T on X.

It is clear from the above remarks that a subsiduary problem is: let (X,T) cover (Y,T), what recursive properties of (Y,T) are "lifted" to (X,T). Some needed results in this direction are also obtained. For

Received July 31, 1964.

[1] Research partially supported under NSF grant NSF-GP-1742.

example: let X, Y be compact, let (X,T) G-cover (Y,T), and let (Y,T) be pointwise almost periodic; then (X,T) is also pointwise almost periodic.

Standing Notation. For the remainder of the paper X will denote a compact metric space, G a complete metric topological group acting on X, ψ a homeomorphism of X such that $xg\psi = x\psi g$ $(x \in X, g \in G)$. Moreover, it is assumed that $Y = X/G$ is compact metric and infinite. Then d will denote the metric on X, Y, and G, π the canonical map of X onto Y, and ϕ the homeomorphism of Y induced by ψ. Thus $(X, \psi$ G-covers (Y, ϕ). (Note: specification of a homeomorphism h of a topological space A onto itself is equivalent to the specification of an action of the integers Z on A. The resulting transformation group is denoted (A, h) rather than (A, Z).)

For $x \in X$, G_x will denote the set $[g \mid g \in G, xg = x]$.

If A is a topological space, h a homeomorphism of A onto A, $x \in A$, then $O(x,h)$ will denote the set $[xh^n \mid n = 0, \pm 1 \cdots]$.

Finally x_0, y_0 will denote fixed elements of X and Y respectively such that $\pi(x_0) = y_0$.

Definition. The group G is *admissible* if the following two conditions are satisfied:

(i) Given open subsets $V_1, \cdots, V_n$ of G there exists an integer p such that $G_{x_0} W_1 \cdots W_p = G$ where $W_i \in [V_1, \cdots, V_n]$ for $1 \leqq i \leqq p$.

(ii) Let f be a continuous function from Y to G and let $\epsilon > 0$. Then there exists $\delta > 0$ such that: if F is a finite subset of Y and u a function from F to G with $d(f(y), u(y)) < \delta$ $(y \in F)$, then there exists a continuous function v from Y to G with $v(y) = u(y)$ $(y \in F)$ and $d(f(y), v(y)) < \epsilon$ $(y \in Y)$.

PROPOSITION 1. *Let G be admissible and let $O(y_0, \phi)$ be dense in Y. Then there exists a homeomophism h of X onto X such that $O(x_0, h)$ is dense and (X, h) covers (Y, ϕ). If, moreover, G is abelian, then (X, h) G-covers (Y, ϕ).*

The proof of Proposition 1 will be broken down into a sequence of lemmas.

Let $C(Y, G)$ denote the set of continuous functions from Y to G provided with the topology of uniform convergence. Then $C(Y, G)$ is a complete metric space with metric $d(u, v) = \sup[d(u(y), v(y)) \mid y \in Y]$.

For $u \in C(Y, G)$ let $\bar{u}: X \to X$ be such that $\bar{u}(x) = x\psi u(\pi(x)\phi)$ $(x \in X)$. Then $\pi(\bar{u}(x)) = \pi(x\psi) = (\pi x)\phi$, $(x \in X)$.

LEMMA 1. *Let $u \in C(Y, G)$. Then*

$$\text{(i)} \qquad \begin{aligned} \bar{u}^n(x) &= x\psi^n u(b\phi) \cdots u(b\phi^n), \quad n \geqq 1 \\ \bar{u}^n(x) &= x\psi^n u(b)^{-1} \cdots u(b\phi^{n+1})^{-1}, \quad n \leqq -1 \end{aligned}$$

where $b = \pi(x)$. (For simplicity $u^n(x)$ will denote $\prod_{i=1}^{n} u(b\phi^i)$ if $n \geqq 1$ and $\prod_{i=1}^{n} u(b\phi^{i-1})^{-1}$ if $n \leqq -1$.)

(ii) *$\bar{u}$ is a homomorphism of X onto X.*

Proof. (i) may be verified directly by induction on n, and (ii) follows from the fact that $\bar{u}$ is continuous and has a continuous inverse given by (i) with $n = -1$.

LEMMA 2. *Let U be open in X,*

$$E(U) = [u/u \in C(Y, G), O(x_0, \bar{u}) \cap U \neq \emptyset].$$

Then $E(U)$ is an everywhere dense open subset of $C(Y, G)$.

Proof. Let $u \in E(U)$. Then there exists n such that $\bar{u}^n(x_0) = x_0\psi^n u^n(x_0) \in U$. For this fixed n, because of the continuity of the various functions involved, there exists $\epsilon > 0$ such that $d(u, v) < \epsilon$ implies $\bar{v}(x_0) \in U$. Thus $E(U)$ is open.

Let $u \in C(Y, G)$, $\epsilon > 0$. Let $\delta > 0$ be as in condition (ii) of the definition of admissibility. Choose V_i $(i = 1, \cdots, n)$ open in G such that diameter $(V_i) < \delta$ $(i = 1, \cdots, n)$ and

$$u(Y) \cup (u(Y))^{-1} \subset \cup [V_i \mid i = 1, \cdots, n].$$

Let p be an integer having the property specified in condition (i) of the definition of admissibility.

Since $O(y_0, \phi)$ is dense in Y and $\pi(U)$ is open, there exists an integer r with $|r| \geqq p$ and $y_0\phi^r \in \pi(U)$. (I assume $r > 0$.) Then there exists $g \in G$ with $x_0\psi^r g \in U$. Let W_i be that element of $[V_1, \cdots, V_n]$ such that $u(y_0\phi^i) \in W_i$ $(i = 1, \cdots, r)$. Since $G_{x_0}W_1 \cdots W_p = G$, $G_{x_0} = G_{x_0\psi^r}$, $r \geqq p$; $G_{x_0\psi^r}W_1 \cdots W_r = G$. Hence there exist $g_i \in W_{ik}$ $(i = 1, \cdots, r)$ with $g \in G_{x_0\psi^r}g_1 \cdots g_r$. Set $F = [y_0\phi^i/i = 1, \cdots, r]$, $w: F \to G$ such that $w(y_0\phi^i) = g_i$ $(i = 1, \cdots, r)$. (Note: w exists since the points $(y_0\phi^i)$ are

distinct.) Finally let v be the continuous extension of w guaranteed by condition (ii). Then $v \in E(U)$ since it extends w and $d(f(y), v(y)) < \epsilon$ $(y \in Y)$. The proof is completed.

We are now in a position to complete the proof of Proposition 1. Let $\mathcal{B}$ be a countable base for the topology on Y, such exists because Y is compact metric. Since $C(Y, G)$ is a Baire space, there exists $u \in \cap [E(U)/U \in \mathcal{B}]$. Then $h = \bar{u}$ is the required homeomorphism.

Since $\bar{u}(x) = x\psi \cdot u(b\phi)$, $\bar{u}(xg) = xg\psi \cdot u(b\phi) = x\psi \cdot gu(b\phi) = x\psi \cdot u(b\phi)g$ $= \bar{u}(x)g$ if G is abelian.

PROPOSITION 2. *Let G be a connected Lie group whose left and right uniform structures coincide. Then G is admissible.*

Proof. Let $V_1, \cdots, V_n$ be open subsets of G. Then there exist $g_1, \cdots, g_n \in G$ and a neighborhood V of the identity with $Vg_i \subset V_i$ $(i = 1, \cdots, n)$. Since the left and right uniform structures coincide, there exists a neighborhood $U \neq \emptyset$ of the identity such that $gUg^{-1} \subset V$ $(g \in G)$. Then $N = \cap [gVg^{-1} \mid g \in G]$ is a non-vacuous neighborhood of the identity with $U \subset N \subset V$ and $gN = Ng$ $(g \in G)$. If $G_{x_0}N^p = G$, then

$$G_{x_0}Nh_1 \cdot Nh_2 \cdots Nh_p = G_{x_0}N^p h_1 \cdots h_p = G$$

where $h_i \in G$. Hence $G_{x_0}N^p = G$ implies that $G_{x_0}W_1 \cdots W_p = G$ for $W_i \in [V_1, \cdots, V_n]$. It thus suffices to show that $G_{x_0}U^p = G$ for some integer p.

We may suppose U compact with $U = U^{-1}$. Let $F = \pi^{-1}\pi(x_0) = x_0G$. Then F is compact and $F = \cup [x_0U^k \mid k = 0, 1, \cdots]$. Note that $G = \cup [U^k \mid k = 0, 1, \cdots]$ since G is connected.)

Since each x_0U^k is compact, there exists an integer r such that x_0U^r has a non-null interior relative to F. Then $F = x_0U^rK$ for some finite subset K of G. Finally, there exists an integer m with $K \subset U^m$. Then $F = x_0U^{r+m}$, which implies that $G = G_{x_0}U^{r+m}$. Thus G satisfies (1).

Let $f \in C(Y, G)$, $\epsilon > 0$. Since $f(Y)$ is compact, there exists $\delta > 0$ such that the ball $B(g, \delta)$ of radius δ about g is contained in a solid space $S(g)$ of diameter $< \epsilon/2$ $(g \in f(Y))$.

Let F be a finite subset of Y, $u\colon F \to G$ with $d(f(y), u(y)) < \delta$ $(y \in F)$. Choose compact disjoint sets U_y $(y \in F)$ such that U_y is a neighborhood of y and $f(U_y) \subset B(f(y), \delta)$ $(y \in F)$. For $y \in F$ let u_y be a continuous function from U_y to $S(f(y))$ such that $u_y(y) = u(y)$ $(y \in F)$ and $u_y(z) = f(z)$ $(z \in \text{boundary of } U_y)$. Finally set $v(z) = u_y(z)$ if $z \in U_y$ and $v(z) = f(z)$ if $z \notin \cup [U_y \mid y \in F]$. Then v is the function required in condition (ii).

In the next few lemmas and propositions (P,T) and (B,T) will denote transformation groups with compact Hausdorff phase spaces P and B and phase group T. Also, f will denote a homomorphism of (P,T) onto (B,T). Thus (P,T) covers (B,T).

LEMMA 3. *Let b be an almost periodic point of (B,T). Then there exists an idempotent u in the enveloping semigroup* [1] *$E(P)$ of (P,T) such that u belongs to a minimal right ideal of $E(P)$ and $xu \in f^{-1}(b)$ $(x \in f^{-1}(b))$.*

Proof. Since b is an almost periodic point of (B,T) there exists an idempotent v in the enveloping semigroup $E(B)$ of (B,T) such that $bv = b$ and v belongs to a minimal right ideal of $E(B)$. [1, Theorem 1]. There exists an idempotent u in $E(P)$ such that u belongs to a minimal right ideal of $E(P)$ and $u\theta = v$; where $\theta\colon E(P) \to E(B)$ is the canonical map induced by f.

Let $x \in f^{-1}(b)$. Then $f(xu) = f(x)(u\theta) = bv = b$; i.e. $xu \in f^{-1}(b)$.

COROLLARY 1. *Let b be an almost periodic point of (B,T). Then there exists $x \in f^{-1}(b)$ such that x is an almost periodic point of (P,T).*

Proof. Corollary 1 follows from Lemma 3 and [1, Lemma 1].

PROPOSITION 3. *Let f be locally one-one, let b be an almost periodic point of (B,T). Then x is an almost periodic point of (P,T) $(x \in f^{-1}(b))$.*

Proof. Let $x \in f^{-1}(b)$, $u \in E(P)$ as in Lemma 3, V a neighborhood of xu such that $f \mid V$ is one-one.

Since $u^2 = u$, $xu \cdot u \in V$ and $xu \in V$. Hence there exists $t \in T$ with xut, $xt \in V$. Now $f(xut) = f(xu)t = bt = f(x)t = f(xt)$. Thus $xut = xt$ since f is one-one on V. This implies that $x = xu$. Consequently x is an almost periodic point of (P,T).

PROPOSITION 4. *Let H be a group acting on P such that (P,T) H-covers (B,T) with canonical map f. Then:* (I) *if b is an almost periodic point of (B,T), x is an almost periodic point of (P,T) $(x \in f^{-1}(b))$.* (II) *If (B,T) is distal, so is (P,T).*

Proof. (I) Let x, b, u be as in Lemma 3. Since $xu \in f^{-1}(b)$, there exists $h \in H$ with $xh = xu$. Moreover, $ht = th$ $(t \in T)$ implies that $uh = hu$. Then $xh = xu = xuu = xhu = xuh$, whence $x = xu$. The proof is completed.

(II). Let w be an idempotent in $E(P)$. It suffices to show that $w = e$ [2, Theorem 1]. Let $x \in P$, $\theta\colon E(P) \to E(B)$ the canonical map. Then $f(xw) = f(x)(w\theta) = f(x)$ since $w\theta$ is an idempotent and $E(B)$ is a group. Hence $xw = xh$ for some $h \in H$. Then $xw = x$ as in the proof of (I). The proof is completed.

PROPOSITION 5. *Let (Y, ϕ) be minimal and distal, G admissible, (xG, G) distal $(x \in X)$, and let $G = G_x K_x$ $(x \in X)$ where K_x is compact. Then there exists a homeomorphism h of X onto X such that (X, h) is minimal and distal.*

Proof. Let h be the homeomorphism constructed in the proof of Proposition 1. I shall show that (X, h) is distal. This will imply that (X, h) is pointwise almost periodic [2, Theorem 1]. Since $O(x_0, h)$ is dense, Proposition 5 will have been proved.

Let $x_1, x_2, x \in X$, (n_i) a sequence of integers such that $x_1 h^{n_i} \to x$ and $x_2 h^{n_i} \to x$. This implies that $y_1 \phi^{n_i} \to y$ and $y_2 \phi^{n_i} \to y$, where $\pi(x_1) = y$, $\pi(x_2) = y_2$, $\pi(x) = y$. Hence $y_1 = y_2$ because (Y, ϕ) is distal. This in turn implies that we may write $x_1 h^{n_i} = x_1 \psi^{n_i} g_i$ and $x_2 h^{n_i} = x_2 \psi^{n_i} g_i$ for some sequence (g_i) of elements of G. (See the construction of h in Proposition 1.)

By the assumptions on G, there exist sequences $(k_i) \subset K_{x_1}$ and $(l_i) \subset K_{x_2}$ such that $x_1 g_i = x_1 k_i$ and $x_2 g_i = x_2 l_i$ for all i. We may assume that $k_i \to k$, $l_i \to l$. Then $x_1 \psi^{n_i} k_i = x_1 k_i \psi^{n_i} = x_1 g_i \psi^{n_i} = x_1 \psi^{n_i} g_i = x_1 h^{n_i} \to x$ implies that $x_1 \psi^{n_i} \to x k^{-1}$ whence $x_1 k \psi^{n_i} \to x$. Similarly $x_2 l \psi^{n_i} \to x$. Now (X, ψ) is distal by Proposition 4. Hence $x_1 k = x_2 l$.

Since $x_1 k_i \to x_1 k$ and $x_2 l_i \to x_2 l$, the sequences $(x_1 g_i)$ and $(x_2 g_i)$ both converge to the same point. Hence $x_1 = x_2$ since $(x_1 G, G)$ is distal. The proof is completed.

COROLLARY 2. *Let (Y, ϕ) be minimal and distal, let G be a connected Lie group whose right and left uniformities coincide. Then there exists a homeomorphism h of X onto X such that (X, h) is minimal and distal.*

Proof. The group G is admissible by Proposition 2.

Let $x \in X$, let V be a compact symmetric neighborhood of the identity of G. Since $\cup[V^n \mid n = 1, \cdots] = G$, $\cup[xV^n \mid n = 1, \cdots] = xG$. Since xG is compact and $xV^n \subset xV^m$ if $m \geqq n$, there exists n with $\operatorname{int}(xV^n) \neq 0$. The compactness of xG implies the existence of a finite subset F of G such that $xV^n = xG$. Set $K = V^n F$. Then K is a compact subset of G with $xK = xG$. Thus $G = G_x K$.

The canonical map of $G \to G \mid G_x = [G_x g \mid g \in G]$ map K onto $G \mid G_x$. Hence G/G_x is compact and the map $g \to xg$ $(g \in G)$ of G onto xG induces

an isomorphism of the transformation group $(G \mid G_x, G)$ onto (xG, G). Since the left and right uniform structures on G coincide, $(G \mid G_x, G)$ is equicontinuous [4; 9.09] hence distal.

Thus G satisfies the hypotheses of Proposition 5. The proof is completed.

An undesirable feature of Proposition 5 is that one must assume (Y, ϕ) distal as well as minimal in order to obtain a minimal set (X, h). This is eliminated in the next proposition.

PROPOSITION 6. *Let (Y, ϕ) be minimal, G admissible and abelian. Then there exists a homeomorphism h of X onto X such that (X, h) is minimal and (X, h) G-covers (Y, ϕ).*

Proof. By Proposition 1 there exists a homeomorphism h of X onto X such that $O(x_0, h)$ is dense and (X, h) G-covers (Y, ϕ). By Proposition 4 (X, h) is pointwise almost periodic. Hence (X, h) is minimal.

Examples. 1. Let (Y, ϕ) be minimal and distal, let G be an admissible group which acts transitively and distally on a compact metric space A. Set $X = A \times Y$, $\psi(a, y) = (a, \phi(y))$ $(a, y)g = (ag, y)$ $(a \in A,\ g \in G,\ y \in Y)$. Then (X, ψ) G-covers (Y, ϕ) and Proposition 5 may be applied to guarantee the existence of a homeomorphism h of X onto X such that (X, h) is minimal and distal.

If we set $Y = S'$ (the circle), $\psi =$ a rotation through one radian, $A = S^n$ (the n-sphere), $G = SO(n+1)$ (the special orthogonal group in $(n+1)$ variables, then the above shows that there exists a minimal, distal discrete flow on $S' \times S^n$.

2. Let Y be a compact manifold, ϕ a diffeomorphism of Y onto Y such that (Y, ϕ) is minimal and distal. Let X be the total space of a differentiable principal fibre bundle over Y with structure group G, a compact connected Lie group. Suppose further that ϕ can be imbedded in a continuous flow on Y. Then it is known that this continuous flow can be lifted to X. Hence there exists a homeomorphism ψ of X onto X such that (X, ψ) G-covers (Y, ϕ) and the preceding propositions may be applied.

3. The Klein bottle K as a minimal set under a discrete flow. (See [3] for a complete discussion of a closely related example.)

For this example R will denote the additive group of real numbers, Z the additive group of integers, and C the additive group of complex numbers. Let $G = R \cdot C$ (semi-direct product) i. e. the underlying space of G is $R \times C$ and $(r, \alpha)(s, \beta) = (r + s, \alpha + e^{\pi i r}\beta)$ $(r, s \in R; \alpha, \beta \in C)$.

The following lemma is immediate.

LEMMA 5. *Let* $f\colon G \to U$ *(the circle group) be such that* $f(r,\alpha) = e^{2\pi ir}$ $(r \in R, a \in C)$. *Then* f *is a homomorphism onto with kernel*

$$L = [(m,\alpha) \mid m \in Z, \alpha \in C].$$

LEMMA 6. *Let* $H = [(m, n + ir) \mid m, n \in Z, r \in R]$. *Then* H *is a closed subgroup of* G *and* $G \mid H = [Hg \mid g \in G]$ *is homeomorphic to* K.

Proof. See [3].

Note that H is also a closed subgroup of L. Let π denote the canonical map of $G \mid H$ onto $G \mid L$.

Now C may be identified with the subgroup $[(0,\alpha) \mid \alpha \in C]$ of G. Hence C acts in a natural way on $G \mid H$.

LEMMA 7. *The group* C *acts transitively on the fibres of* $G \mid H$ *over* $G \mid L$. *(Thus* $(G \mid H) \mid C = G \mid L$.)

Proof. It suffices to show that $HgC = HgL$ $(g \in G)$. Since $C \subset L$, $HgC \subset HgL$.

Let $l = (m,\alpha) \in L$, $g = (r,\beta) \in G$. We must find $h \in H$ and $c \in C$ with $gl = hgc$. Set $h = (m, 0)$ and $c = (0,\gamma)$. Then $gl = (r + m, \beta + e^{\pi ir}\alpha)$, $hg = (r + m, e^{\pi im}\beta)$, $hgc = (r + m, e^{\pi im}\beta + e^{\pi i(r+m)}\gamma)$. Then the equation $e^{\pi im}\beta + e^{\pi i(r+m)}\gamma = \beta + e^{\pi ir}\alpha$ determines γ.

LEMMA 7. *Let* $t \in R$, $\lambda\colon C \to C$ *such that* $\lambda(0,\alpha) = (0, e^{-\pi it}\alpha)$ $(\alpha \in C)$. *Then* $gc(t,0) = g(t,0)\lambda(c)$ $(g \in G, c \in C)$.

Proof. Verify directly.

LEMMA 8. *Let* $t \in R$, $\psi\colon G \mid H \to G \mid H$ *such that* $\psi(Hg) = Hg(t,0)$ $(g \in G)$. *Then* ψ *induces a map* $\phi\colon G \mid L \to G \mid L$ *such that* ϕ *is a rotation through* $2\pi t$ *radians when* $G \mid L$ *is identified with* U.

Proof. By Lemma 7, ψ is fibre preserving. Hence ϕ exists. Moreover $\phi(Lg) = Lg(t,0)$ $(g \in G)$. If $g = (r,\alpha)$ then f identifies Lg with $e^{2\pi ir}$ and $Lg(t,0)$ with $e^{2\pi i(r+t)} = e^{2\pi ir}e^{2\pi it}$.

Now fix $t \in R$ such that the resulting transformation group (U,ϕ) is minimal and distal. Then (K,ψ) covers (U,ϕ). However it doesn't quite C-cover (u,ϕ) because $kc\psi = k\psi\lambda(c)$ $(k \in K, c \in C)$ where $\lambda(c) = e^{-\pi it}c$ $(c \in C)$. Thus the preceding propositions cannot be applied directly to this situation. However, the proofs may be modified slightly to yield a minimal distal discrete flow on K.

LEMMA 9. *Let (K,ρ) cover (U,ϕ) and let $kc\rho = k\rho\lambda(c)$ $(k \in K, c \in C)$. Then (K,ρ) is distal.*

Proof. Let $k \in K$, u an idempotent in the enveloping semigroup E of (K,ρ). I must show that $ku = k$.

Since (U,ϕ) is distal and (K,ρ) covers (U,ϕ), ku is in the same fibre as k. Hence $ku = kc$ for some $c \in C$.

It is clear that $ldp = lp\lambda(d)$ $(l \in K,\ p \in E,\ d \in C)$. Hence $ku = kuu = kcu = ku\lambda(c)$. Set $l = ku$. Then $l = l\lambda(c)$. Thus $lp = l\lambda(c)p = lp\lambda^2(c)$ $(p \in E)$. Hence $\lambda^2(c) \in \cap\,[C_{lp} \mid p \in E]$.

Since (U,ϕ) is minimal, the set lE intersects every fibre. Since C is abelian $C_x = C_y$ if x and y are in the same fibre. Consequently $\lambda^2(c) \in C_x$ $(x \in K)$. This implies $\lambda^2(c) = 0$; whence $c = 0$ and $kc = k$.

Now let u be a continuous function from U to C and set $\bar{u}(k) = k\psi u((\pi k)\phi)$ as in Proposition one. Then
$\bar{u}(kc) = (kc)\psi u(\pi k\phi) = k\psi\lambda(c)u(\pi k\phi) = k\psi u(\pi k\phi)\lambda(c) = \bar{u}(k)\lambda(c)$ $(k \in K, c \in C)$ since C is abelian. Thus $(K,\bar{u})$ is distal by Lemma 9. Hence in order to product a minimal, distal, discrete flow on K we must find u such that $O(k_0,\bar{u})$ is dense for some k_0 in K. The proof that this can be done is so similiar to the proof of Proposition 1 that I shall merely sketch it.

One verifies directly that if u is a continuous function from U to C, then $\bar{u}$ defined above is a homeomorphism of K onto K such that

$$\bar{u}^n(k) = k\psi^n\Big[\sum_{i=0}^{n-1}\lambda^i(u(\pi k\phi^{i+1}))\Big].$$

Let V be open in K, $k_0 = \{H\} \in K$, $E(V) = [u \mid O(k_0,\bar{u}) \cap V \neq \emptyset]$. Then it suffices to show that $E(V)$ is open and everywhere dense in the space of continuous functions from U to C provided with the topology of uniform convergence.

That $E(V)$ is open is immediate The proof that it is dense will be broken up into a sequence of lemas.

Let $\alpha = \alpha_1 + i\alpha_2 \in C$ and $a \in R$ with $a > 0$. Then $Sq(\alpha,a)$ will denote the set $[x \mid x = x_1 + ix_2 \in C, |x_1 - \alpha_1| \leqq a, |x_2 - \alpha_2| \leqq a]$, and $S(\alpha,a)$ will denote the set $[x \mid |\alpha - x| \leqq a]$.

LEMMA 10. *Let $\beta \in C$. Then $kSq(\beta,2) = kC$ $(k \in K)$.*

Proof. Let $k = Hg$, $q = (r,\alpha)$, $c = c_1 + ic_2 \in C$. Then $kc = Hgc$, $gc = (r,\alpha)(0,c) = (r,\alpha + e^{\pi ir}c)$. To find $h \in H$, $x \in Sq(\beta,2)$ such that $gc = hgx$. Let $h = (0,n + is)$, $x = x_1 + ix_2$. Then

$$hgx = (r, n + is + e^{\pi ir}x + \alpha).$$

Thus we must have $n+is+e^{\pi i r}x=e^{\pi i r}c$. Since s may be chosen arbitrarily, we need only consider the real part of the above equation.

The equation to be satisfied becomes:

$$n+x_1\cos\pi r-x_2\sin\pi r=c_1\cos\pi r-c_2\sin\pi r.$$

Since $\cos^2\pi r+\sin^2\pi r=1$, $\max(|\cos\pi r|,|\sin\pi r|)\geqq\frac{1}{2}$. Let us suppose that $|\cos\pi r|\geqq\frac{1}{2}$. Then set $x_2=\beta_2$, $x_1=\gamma_1+\beta_1$. We must be able to find $n\in Z$ and $0\leqq\gamma_1\leqq 2$ such that:

$$n+\gamma_1\cos\pi r=c_1\cos\pi r-c_2\sin\pi r+\beta_2\sin\pi r-\beta_1\cos\pi r.$$

This is always possible since the quantity $\gamma_1\cos\pi r$ covers an interval of length one as γ_1 varies between 0 and 2.

LEMMA 11. *Let $\alpha,\beta\in C$, $a,b\in R$ with $a,b>0$. Then:* (1) $\lambda S(\alpha,a)=S(\lambda(\alpha),a)$ (ii) $Sq(\alpha,a)+Sq(\beta,b)\supset Sq(\alpha+\beta,a+b)$ (iii) $S(\alpha,a)\subset Sq(\alpha,a)$ (iv) $S(\alpha,a)\supset Sq(\alpha,a/\sqrt{2})$.

Proof. Immediate.

I shall now complete the proof that $E(V)$ is dense. Let w be a continuous function from U to C and let $\epsilon>0$. Let $a\in R$ with $0<a<\epsilon/2$ and $\alpha_1,\cdots,\alpha_n\in C$ with $w(U)\subset\cup[S(\alpha_i,a)\mid i=1,\cdots,n]$. Let $p\in Z$ with $p\dfrac{a}{\sqrt{2}}\geqq 2$, $r\in Z$ with $r\geqq p$ and $(\pi k_0)\phi^r\in\pi(V)$. For $i=1,\cdots,r$ let $\beta_i\in[\alpha_1,\cdots,\alpha_n]$ be such that $w((\pi k_0)\phi^i)\in S(\beta_i,a)$. Then $\lambda^{i-1}S(\beta_i,a)=S(\lambda^{i-1}\beta_i,a)=S(\gamma_i,a)$ where $\gamma_i=\lambda^{i-1}\beta_i$, $i=1,\cdots,r$. Now
$S(\gamma_r,a)+S(\gamma_{r-1},a)+\cdots+S(\gamma_1,a)$
$\supset Sq(\gamma_r,a/\sqrt{2})+\cdots+Sq(\gamma_1,a/\sqrt{2})\supset Sq(\sum\gamma_i,ra/\sqrt{a})\supset Sq(\sum\gamma_i,2)$.
Since C is transitive on fibres, there exists $c\in C$ with $k_0\psi^rc\in V$. By Lemma 10 we may assume $c\in Sq(\sum\gamma_i,2)$. Let $c_i\in S(\beta_i,a)$ be such that $\sum\lambda^{i-1}c_i=c$.

It is clear that there exists a continuous function u from U to C such that $u(\pi k\cdot\phi^i)=c_i$ $(i=1,\cdots,r)$ and $d(w(y),u(y))<\epsilon$ $(y\in U)$.

Then $d(w,u)<\epsilon$ and $u\in E(V)$ because

$$\bar{u}^r(k_0)=k_0\psi^r[\sum_{i=0}^{r-1}\lambda^i(u(\pi k_0\phi^{i+1}))]=k_0\psi^r(\sum_{i=0}^{r-1}\lambda^i(c_{i+1}))=k_0\psi^rc\in V.$$

4. If in the preceding example we set $\psi_s(Hg)=Hg(st;0)$ $(g\in G)$ and $u_s(y)=su(y)$ $(0\leqq s\leqq 1)$, then $\bar{u}_s(k)=k\psi_su_s(\pi k\phi)$ sets up a homotopy between $\bar{u}$ and the identity. Thus $\bar{u}$ can be lifted to a continuous mapping v of the torus T onto itself. It is clear that v is a homeomorphism of T onto

itself such that (T, v) is minimal and distal. Let σ be a homeomorphism of T onto itself of period 2 such that $v\sigma = \sigma v$. Set $w = v\sigma$. Then (T, w) is also minimal and distal. Since either w or v is orientation reversing, this shows that there exists an orientation reversing homeomorphism of T onto itself under which T is minimal and distal.

WESLEYAN UNIVERSITY.

REFERENCES.

[1] Robert Ellis, "A semigroup associated with a transformation group," *Transactions American Mathematical Society*, vol. 94 (1960), pp. 272-281.

[2] ———, "Distal transformation groups," *Pacific Journal of Mathematics*, vol. 8 (1958), pp. 401-405.

[3] ———, "Global sections of transformation groups," to appear in *Illinois Journal of Mathematics*.

[4] W. H. Gottschalk and G. A. Hedlund, *Topological dynamics*, American Mathematical Society Colloquium Publications, vol. 36, Providence, 1955.

THE CONJUGATE LOCUS OF A RIEMANNIAN MANIFOLD.

By Frank W. Warner.*

Introduction. The conjugate locus (considered as a subset of the tangent space to a point of a Riemannian manifold) splits naturally into two subsets—the regular locus and the singular locus; the latter, roughly speaking, consists of those points which occur at intersections in the conjugate locus. We describe in this paper the regular conjugate locus and the nature of the exponential map nearby. Our main results are these: The regular conjugate locus is dense in the conjugate locus and is a submanifold of the tangent space of codimension one. If the order of a regular conjugate point, i. e., the dimension of the kernel of the differential of the exponential map there, is greater than or equal to 2, the kernel must actually be tangent to the regular conjugate locus. This was proved for analytic Finsler spaces by J. H. C. Whitehead in [11] in the case that the order of the conjugate point is greater than half the dimension of the manifold. We obtain normal forms for the exponential map on neighborhoods of all regular conjugate points except for certain of the order 1 cases. As a corollary we obtain a new proof of the result of Morse and Littauer [5] (for analytic Finsler spaces) and of Savage [8] (for the C^∞ case) that the exponential map is never 1:1 on any neighborhood of any conjugate point.

Our results do not depend on a Riemannian structure and hold for what we call regular exponential maps. These are maps of a tangent space into a manifold satisfying certain conditions on their first and second order differentials and a continuity property for their singular points which we call conjugate points. We define and develop some of the properties of regular exponential maps in Section 2 after a preliminary Section 1 where we establish our notational conventions. Our main theorems on the regular conjugate locus of a regular exponential map are in Section 3. Section 4 contains enough of the elementary geometry of Finsler spaces to prove that the exponential maps of a Finsler space are regular. Further comparison is given there between our results and techniques and those of Whitehead, Morse,

Received July 7, 1964.

* This research was begun while the author held a National Science Foundation Cooperative Graduate Fellowship at M. I. T. It was completed with further support from the National Science Foundation.

Littauer and Savage mentioned above. Finally in Section 5 we prove, both as a corollary of the section on Finsler spaces and directly from basic facts of Riemannian geometry, that the exponential maps of a Riemannian manifold are regular.

I am indebted to I. M. Singer for suggesting this problem to me and for the many fruitful ideas I obtained from conversations with him.

1. **Preliminaries.** We need to fix some notation and conventions to be used throughout the paper. Manifolds will be locally euclidean, second countable, Hausdorff spaces with differentiable structure. All manifolds, maps, vector fields, etc., will be differentiable of class C^∞. By a *submanifold* of a manifold M is meant a manifold N together with a $1:1$ immersion of N into M. Let m be a point of a manifold M. We denote the linear space of k-th order tangent vectors to M at m by $M^k{}_m$ (cf. [2] for the definitions of higher order tangent vectors and differentials in the form we will use them). The first order tangent space $M^1{}_m$ is denoted simply by M_m. M_m is naturally a manifold since it is a linear space over the reals; and, with the notation we have adopted, the k-th order tangent space to M_m at a point p in M_m is denoted by $(M_m)^k{}_p$. $T(M)$ is the tangent bundle of M and $T^*(M)$ the dual bundle. If $f\colon M \rightarrow N$ is a differentiable map of manifolds, then df will denote its first order differential and δf the dual of df. We call a point m in M a *singularity* of f of *order* or *multiplicity* $k > 0$ if the dimension of the kernel of $df \mid M_m$ is k. The k-th order differential of f which maps $M^k{}_m$ into $N^k{}_{f(m)}$ will be denoted by d^kf for $k > 1$. If ω is a 1-form on M, then (to eliminate the usual confusion of the factor of $(\frac{1}{2})$) we will simply define $d\omega$ to be the 2 form such that:

$$d\omega(X,Y) = (\tfrac{1}{2})\{X(\omega(Y)) - Y(\omega(X)) - \omega[X,Y]\}$$

for all vector fields X and Y on M. If Y is a vector field on M than $Y(m)$ is its value at m. R will denote the real line. The tangent vector to a curve $\sigma\colon R \rightarrow M$ will be denoted by $\sigma_*(t)$, i.e., $\sigma_*(t) = d\sigma(d/dt(t))$. If Y is a vector field along a curve σ then $Y(t) = Y(\sigma(t))$. If f is a function on R (all functions will be real valued) then we let $f'(t)$ denote its derivative at t. By a *ray* r in the tangent space M_m is meant any non-trivial linear map r of R into M_m. Let p be a point of M_m. By *the ray in* M_m *through* p is meant the mapping of the *positive* real numbers $R^+ \rightarrow M_m$ sending t to tp. Finally, neighborhoods will always be open.

We will use later the following form of the well known implicit function theorem:

THEOREM 1.1. *Let N and M be manifolds of dimensions c and d respectively. Let $f: N \to M$ be C^∞ and let $m \in M$. If df is a surjective map of $N_n \to M_m$ for every n in $f^{-1}(m)$ and if $T = f^{-1}(m)$, then T can be given a manifold structure of dimension $(c-d)$ such that the inclusion $i: T \to N$ is a submanifold with the relative topology.*

We use freely throughout the natural isomorphism of $M_m \# M_m$ with $M^2{}_m/M_m$, where $\#$ denotes the symmetric product. If V_1 and V_2 are subspaces of a vector space V then $V_1 \# V_2$ is the subspace of $V \# V$ generated by elements of the form $v_1 \# v_2$ for $v_1 \in V_1$ and $v_2 \in V_2$. If the dimensions of V_1, V_2 and $V_1 \cap V_2$ are a, b and c, respectively, then the dimension of $V_1 \# V_2$ is $(ab - c(c-1)/2)$.

2. Regular exponential maps. Throughout this section and the next, M will be a fixed d-dimensional manifold, and m a fixed point in M. If F is a smooth map of M_m into M and $p \neq 0$ a point of M_m, we let $N(p)$ denote the null space of dF at p and r_p denote the tangent space at p to the ray through p. The second order differential d^2F of F induces a linear transformation, which we also denote by d^2F, from $(M_m)^2{}_p/(M_m)_p$ to $M^2{}_{F(p)}/dF((M_m)_p)$. Under this mapping the subspace $r_p \# N(p)$ is mapped into $M_{F(p)}/dF((M_m)_p)$.

Definition. A map $F: M_m \to M$ is called a *regular exponential map* if the following three conditions are satisfied:

(R1) F is C^∞ on M_m except possibly at the origin where it is at least C^1, and $dF(r_*(t)) \neq 0$ for all t, for every ray r in M_m.

(R2) The subspace $r_p \# N(p)$ of $(M_m)^2{}_p/(M_m)_p$ is mapped isomorphically onto $M_{F(p)}/dF((M_m)_p)$ by d^2F for each $p \neq 0$ in M_m.

(R3) For each non zero point p in M_m there exists a convex neighborhood U of p such that the number of singularities of F (counted with multiplicities) on $r \cap U$, for each ray r which intersects U, is constant and equals the order of p as a singularity of F.

We will see in Sections 4 and 5 that the exponential map for a Riemannian manifold, and more generally for a Finsler space, is, when restricted to a particular tangent space M_m, a regular exponential map. In the Riemannian case the property (R1) holds since the exponential map sends the rays in M_m into the non-trivial geodesics through m, and the tangent vector to a non-trivial geodesic can never vanish. It will be shown that property (R2) is equivalent to the fact that if one considers the Jacobi fields along a

geodesic through m with initial value zero at m, then at any point n on this geodesic the full tangent space M_* is spanned by the values of these Jacobi fields at n and the covariant derivatives of those Jacobi fields which vanish at n. Property (R3) is the well known continuity property for conjugate points considered with their multiplicities.

We first investigate the property (R2) and the way d^2F acts on $(M_m)^2{}_p/(M_m)_p$. Suppose F is a mapping of M_m into M satisfying (R1). Let p be a non-zero point of M_m and $\sigma(t)$ the ray in M_m through p. Let $A(t)$ be a C^∞ vector field along σ such that $A(1) \in N(p)$. Let $e_1(t), \cdots, e_d(t)$ be C^∞ vector fields along $F \circ \sigma$ which span $M_{F(\sigma(t))}$ for all t in the domain of σ. Let $Y(t) = dF(A(t))$, a C^∞ vector field along $F \circ \sigma$. Then $Y(t) = \sum_{i=1}^{d} f_i(t) e_i(t)$, where the f_i are C^∞ and $Y(1) = 0$. We define

$$Y'(1) = \sum_{i=1}^{d} f_i'(1) e_i(1).$$

It is easily checked that since $Y(1) = 0$, $Y'(1)$ is a well defined element of $M_{F(p)}$ which does not depend on the basis $e_1(t), \cdots, e_d(t)$. In case M is affinely connected, $Y'(1)$ is the same as the covariant derivative of Y along $F \circ \sigma$ with respect to $(F \circ \sigma)_*$ at $F(\sigma(1))$. Extend σ_* and A from σ to vector fields on a neighborhood of p. Then $(\sigma_* A)(p) \in (M_m)^2{}_p$ and $\sigma_*(1) \# A(1) = (\sigma_* A)(p) + (M_m)_p$ as elements of $(M_m)^2{}_p/(M_m)_p$.

LEMMA 2.1.

$d^2F(\sigma_*(1) \# A(1)) = d^2F((\sigma_* A)(p) + (M_m)_p) = Y'(1) + dF((M_m)_p)$ *and in fact* $d^2F((\sigma_* A)(p)) = Y'(1)$. *If* $A(1) = 0$, *then* $Y'(1) \in dF((M_m)_p)$. *If* $A(1) \neq 0$ *and if* F *satisfies* (R2), *then* $Y'(1) \notin dF((M_m)_p)$ *and, in particular,* $Y'(1) \neq 0$.

Proof. The first equality is obvious and the second follows from the fact $d^2F((\sigma_* A)(p)) = Y'(1)$. A simple computation shows that if g is any C^∞ function on a neighborhood of $F(p)$, then $d^2F((\sigma_* A)(p))(g) = Y'(1)(g)$ which proves the desired equality. If $A(1) = 0$, then $\sigma_*(1) \# A(1) = 0$, so $d^2F(\sigma_*(1) \# A(1)) = 0$ as an element of $M^2{}_{F(p)}/dF((M_m)_p)$. Hence $Y'(1) \in dF((M_m)_p)$. If, however, $A(1) \neq 0$, then $\sigma_*(1) \# A(1)$ is a non-zero element of $r_p \# N(p)$. If F satisfies (R2), d^2F is one-one on $r_p \# N(p)$ so $d^2F(\sigma_*(1) \# A(1)) \neq 0$. Hence $Y'(1) \notin dF((M_m)_p)$, and, in particular, $Y'(1) \neq 0$. q. e. d.

Before proceeding with regular exponential maps we need some information about square roots of C^∞ functions. Suppose f is a non-negative C^∞

function on the real line; i.e., $f(t) \geqq 0$ all t. It can be shown that f has a C^1 square root whose second derivatives exist everywhere but that, in general, f does not have a C^2 square root. We need the following special case in which f has a C^∞ square root. A zero of f will be called *finite* if some derivative of f does not vanish at that point. The *order* of a finite zero is n if the n-th derivative is the first non-vanishing one.

LEMMA 2.2. *If f is a real valued, non-negative, C^∞ function on the real line all of whose zeros are of finite order, then f has a C^∞ square root.*

Proof. Suppose f has a finite zero of order n; and, for convenience, assume it occurs at 0. Then there exists a neighborhood of 0 on which f is strictly positive except at 0. On this neighborhood there exists a C^∞ function g such that $f(t) = t^n g(t)$, where $(n!)g(0) = (d^n f/dt^n)(0) \neq 0$. Since $f(t) \geqq 0$, n must be even and g strictly positive on this neighborhood. Let $n = 2m$. Then on this neighborhood f has a C^∞ square root; namely, $f^{\frac{1}{2}}(t) = t^m g^{\frac{1}{2}}(t)$. The square root changes sign at 0 if and only if m is odd. One obtains a C^∞ square root of f on the real line by choosing a square root at a point where f is non zero. The choice of square root at every other point is then determined by the requirement that $f^{\frac{1}{2}}$ change sign only at a zero of f of order n when $n/2$ is odd. q. e. d.

In the following sequence of lemmas we establish the existence of a useful coordinate system along the image under F of a ray in M_m.

LEMMA 2.3. *Let $F\colon M_m \to M$ satisfy* (R1) *and* (R2). *Let p be a non-zero point in M_m and σ the ray in M_m through p. Let $A(t)$ be a C^∞ non-vanishing vector field along σ, and let $Y(t) = dF(A(t))$, a C^∞ vector field along $F\circ\sigma$. Then $Y(t) = f(t)\cdot e(t)$, where $e(t)$ is a non-vanishing C^∞ vector field along $F\circ\sigma$, $f(t)$ is a C^∞ function, and $f' \neq 0$ whenever $f = 0$.*

Proof. Let $e_1(t),\cdots,e_d(t)$ be smooth vector fields along $F\circ\sigma$ forming a basis of $M_{F(\sigma(t))}$ for all t in the domain of σ. Then $Y(t) = \sum_{i=1}^{d} f_i(t)e_i(t)$, where the f_i are C^∞. Suppose $Y(t_0) = 0$. Then $dF(A(t_0)) = 0$. By Lemma 2.1, $Y'(t_0) \neq 0$. Hence, not all of the $f_i'(t_0)$ can be zero whenever t_0 is a zero of Y. This implies, in particular, that the zero's of Y are isolated.

An easy computation shows that $\sum_{i=1}^{d} f_i^2(t)$ is a non-negative C^∞ function whose zeros are all of second order. So by Lemma 2.2, this function has a

C^∞ square root, which we denote by $f(t)$. The zeros of $f(t)$ are isolated and coincide with the zeros of $Y(t)$, and are all of order one. Define:

$$
\begin{aligned}
e(t) &= \frac{Y(t)}{f(t)} = \sum_{i=1}^{d} \frac{f_i(t)}{f(t)} e_i(t), \text{ if } Y(t) \neq 0, \text{ and} \\
e(t) &= \sum_{i=1}^{d} \frac{f_i'(t)}{f'(t)} e_i(t), \text{ if } Y(t) = 0.
\end{aligned} \tag{2.1}
$$

Then $e(t)$ is a non-vanishing vector field along $F \circ \sigma$ such that $Y(t) = f(t)e(t)$. That $e(t)$ is C^∞ follows immediately from the fact that if t_0 is a zero of Y, then on a neighborhood of t_0, $f_i(t) = (t - t_0)k_i(t)$ and $f(t) = (t - t_0)g(t)$, where the $k_i(t)$ and $g(t)$ are C^∞, $k_i(t_0) = f_i'(t_0)$, and $g(t_0) = f'(t_0) \neq 0$.

q. e. d.

LEMMA 2.4. *Let $\gamma: (a,b) \to M: t \to \gamma(t)$ be a one-one C^∞ curve in M with $\gamma_*(t) \neq 0$, $t \in (a,b)$. Let $Y_1, \cdots, Y_d$ be C^∞ vector fields along γ which span $M_{\gamma(t)}$ for all $t \in (a,b)$. Let $t_0 \in (a,b)$. Then there exists a coordinate system $y_1, \cdots, y_d$ on a neighborhood V of $\gamma(t_0)$ such that $\partial/\partial y_i(\gamma(t)) = Y_i(\gamma(t))$ for $t \in \gamma^{-1}(V)$ and for $i = 1, \cdots, d$.*

Proof. There exists a cubic coordinate system $x_1, \cdots, x_d$ on a neighborhood U of $\gamma(t_0)$ and a neighborhood, say (a,b) of t_0, such that $\gamma(a,b)$ is the slice of U given by $x_2 = \cdots = x_d = 0$, and $x_1(\gamma(t)) = t$. Along γ

$$\partial/\partial x_i(\gamma(t)) = \sum_{j=1}^{d} a_{ij}(t) Y_j(t)$$

where the a_{ij} are C^∞ functions on (a,b) and the matrix $A(t) = (a_{ij}(t))$ is non-singular for $t \in (a,b)$. Define

$$y_i(p) = \int_a^{x_1(p)} a_{1i}(t)\,dt + \sum_{k=2}^{d} (a_{ki} \circ x_1(p))(e^{x_k(p)} - 1)$$

for $i = 1, \cdots, d$, and for $p \in U$. The $y_1, \cdots, y_d$ are then C^∞ functions on U such that along γ, $(\partial y_i/\partial x_j)(\gamma(t)) = a_{ji}(t)$ for $i, j \in \{1, \cdots, d\}$. Since

$$(dy_1, \cdots, dy_d) = (dx_1, \cdots, dx_d) \cdot A(t_0)$$

the dy_i are independent at $\gamma(t_0)$ and hence the y_i form a coordinate system on a neighborhood $V \subset U$ of $\gamma(t_0)$. At $\gamma(t)$

$$
\begin{aligned}
(\partial/\partial x_1, \cdots, \partial/\partial x_d) &= (Y_1, \cdots, Y_d) \cdot A^*(t) \\
&= (\partial/\partial y_1, \cdots, \partial/\partial y_d) \cdot A^*(t)
\end{aligned}
$$

where $*$ denotes transpose; so since $A(t)$ is non singular,

$$(\partial/\partial y_1(\gamma(t)), \cdots, \partial/\partial y_d(\gamma(t))) = (Y_1(\gamma(t)), \cdots, Y_d(\gamma(t))).$$

q. e. d.

LEMMA 2.5. *Let $F: M_m \to M$ satisfy* (R1) *and* (R2). *Let $p \in M_m$ be a singularity of order k for F. Let σ be the ray in M_m through p. Then there exist coordinate systems $x_1, \cdots, x_d$ and $y_1, \cdots, y_d$ on neighborhoods U and V of p and $F(p)$ respectively, such that $F(U) \subset V$ and for $t \in \sigma^{-1}(U)$,*

$$dF(\partial/\partial x_j(\sigma(t))) = f_j(t)\,(\partial/\partial y_j(F(\sigma(t))))$$

for $j = 1, \cdots, d$, where the f_j are C^∞ functions such that $f_j(t) \neq 0$ for $j = 1, \cdots, d-k$. For $j = d-k+1, \cdots, d$, $f_j(t)$ is zero only at $t = 1$, and there $f_j'(1) > 0$.

Proof. Choose a basis $A_1, \cdots, A_d$ of $(M_m)_p$ such that $A_{d-k+1}, \cdots, A_d$ span the null space of $dF(p)$, and extend to C^∞ vector fields $A_1(t), \cdots, A_d(t)$ along σ which span $(M_m)_{\sigma(t)}$ for each t. We consider t only in some interval containing 1 where $F \circ \sigma$ is a 1:1 curve. Let $Y_j(t) = dF(A_j(t))$ for $j = 1, \cdots, d$. By Lemma 2.3, $Y_j(t) = f_j(t) e_j(t)$ where $e_j(t)$ is a non-vanishing C^∞ vector field along $F \circ \sigma$ and $f_j(t)$ is a C^∞ function such that $f_j'(t) \neq 0$ whenever $f_j(t) = 0$, for $j = 1, \cdots, d$. Now

$$f_1(1) \neq 0, \cdots, f_{d-k}(1) \neq 0, \text{ and } f_{d-k+1}(1) = \cdots = f_d(1) = 0.$$

By changing the signs of e_j and f_j, if necessary, we can arrange that $f_{d-k+1}'(1) > 0, \cdots, f_d'(1) > 0$. There exists an interval $I \subset R$ containing 1 on which the only zero of an f_j occurs at 1. We claim that $e_1(1), \cdots, e_d(1)$ are linearly independent vectors in $M_{F(p)}$. Indeed, $e_1(1), \cdots, e_{d-k}(1)$ span the range of $dF(p)$ in $M_{F(p)}$, and by (2.1) and Lemma 2.1,

$$d^2F(\sigma_*(1) \# A_j(1)) = e_j(1)/c_j + dF((M_m)_p)$$

for $j = d-k+1, \cdots, d$ where c_j is a non zero constant. Since the elements $\sigma_*(1) \# A_j(1)$ for $j = d-k+1, \cdots, d$ form a basis of $r_p \# N(p)$, the corresponding $e_j(1) + dF((M_m)_p)$ form a basis of $M_{F(p)}/dF((M_m)_p)$ by (R2). Hence $e_1(1), \cdots, e_d(1)$ is a basis of $M_{F(p)}$. Therefore, since the $e_i(t)$ are C^∞ vector fields along $F \circ \sigma$, $e_1(t), \cdots, e_d(t)$ span $M_{F(\sigma(t))}$ for t in some neighborhood of 1 in R. By Lemma 2.4, there exists a coordinate system $x_1, \cdots, x_d$ on a neighborhood U of p in M_m such that $\partial/\partial x_j(\sigma(t)) = A_j(t)$ for all j, and a coordinate system $y_1, \cdots, y_d$ on a neighborhood V of $F(p)$ such that $\partial/\partial y_j(F(\sigma(t))) = e_j(t)$ for all j. U and V can be chosen so that $F(U) \subset V$ and $\sigma^{-1}(U) \subset I$. These are the desired coordinate systems.

q. e. d.

3. The regular conjugate locus.

Definition. The singularities of a regular exponential map $F: M_m \to M$

will be called *conjugate points*. The *order* of a conjugate point is its order as a singularity of F. The set of all conjugate points in M_m is called the *conjugate locus* and is denoted by $C(m)$. A conjugate point $p \in M_m$ is called *regular* if there exists a neighborhood U of p such that each ray of M_m contains at most one point in U which is a conjugate point. A conjugate point which is not regular is called a *singular* or *intersection* point.

We let $C^R(m)$ denote the regular points in $C(m)$ and $C^S(m)$ denote the intersection points. If U is an open set in M_m, we set $C_U(m) = C(m) \cap U$, $C_U{}^R(m) = C^R(m) \cap U$ and $C_U{}^S(m) = C^S(m) \cap U$. Note that, by (R3), $C^R(m)$ contains all the 1st order conjugate points.

Our ultimate objective is to determine the nature of the singularities of a regular exponential map and to give a decomposition of the conjugate locus into subsets admitting particularly nice descriptions. In the following theorems we carry out this program for the regular conjugate locus.

THEOREM 3.1. *Let $F: M_m \to M$ be a regular exponential map. Then the regular conjugate locus $C^R(m)$ is an open everywhere dense subset of $C(m)$ which can be given a manifold structure of dimension $d-1$ such that the inclusion $i: C^R(m) \to M_m$ is a submanifold with the relative topology, and such that $(M_m)_p = (C^R(m))_p \oplus r_p$ for every $p \in C^R(m)$.*

Proof. That $C^R(m)$ is open in $C(m)$ follows immediately from the definition of regular conjugate points. To prove $C^R(m)$ is dense in $C(m)$ we show that $C^S(m)$ is nowhere dense in $C(m)$. Since $C^S(m)$ is closed in $C(m)$, we must show that the interior of $C^S(m)$ is empty. Suppose, on the contrary, that there exists a point $p \in \operatorname{int}(C^S(m))$, and suppose the rank of dF at p is $d-k$. Then there exists a neighborhood U of p in M_m such that $C_U(m) \subseteq C^S(m)$; and, by (R3), U can be chosen so small that for each ray r in M_m which intersects U, the number of conjugate points (counted with multiplications) on $r \cap U$ is k. Since $p \in C^S(m)$, there exists a ray which intersects U in at least two conjugate points, at each of which necessarily the rank of dF is greater than or equal to $d-k+1$. Choose one of these points; call it p_1. Then $p_1 \in \operatorname{int}(C^S(m))$; and we repeat the above for p_1. In a finite number of such steps we obtain a point $p_i \in \operatorname{int}(C^S(m))$, $1 \leqq i \leqq k-1$, such that rank $(dF(p_i)) = d-1$. But then $p_i \in C^R(m)$ which is a contradiction. Hence $C^R(m)$ is dense in $C(m)$.

Finally we show that, with the relative topology, $C^R(m)$ has a manifold structure of dimension $d-1$ such that the inclusion $i: C^R(m) \to M_m$ is a submanifold. Let $p \in C^R(m)$. It suffices to prove that there exists a neighborhood W of p in M_m such that $C_W(m) \subseteq C_W{}^R(m)$ and such that

$C_W(m)$ has a manifold structure of dimension $d-1$ such that $i\colon C_W(m)\to M_m$ is a submanifold in the relative topology.

Suppose the order of p is k. Choose coordinate neighborhoods U, $x_1,\cdots,x_d$ of p and V, $y_1,\cdots,y_d$ of $F(p)$ as in Lemma 2.5. With respect to these coordinate systems, $dF\mid U$ is a matrix with entries C^∞ functions on U. Let Δ_i be the i-th elementary symmetric function in the eigenvalues of dF for $i=0,\cdots,d-1$. The Δ_i are C^∞ functions on U. Let σ be the ray in M_m through p. Then along σ, dF is diagonal with entries $f_1(t),\cdots,f_d(t)$, C^∞ functions of t. Moreover, the only zeros of the f_i on $\sigma^{-1}(U)$ occur at $t=1$, where $f_j(1)=0$ for $j=d-k+1,\cdots,d$. And $f_{d-k+1}'(1)>0,\cdots,f_d'(1)>0$. Hence $\Delta_{k-1}(p)=0$ and

$$\sigma_*(1)(\Delta_{k-1})=\Big(\sum_{j=d-k+1}^{d}f_j'(1)\Big)(f_1(1)\cdot\cdots\cdot f_{d-k}(1))$$

since $\Delta_{k-1}(\sigma(t))$ is the sum of the products $f_1(t)\cdot\cdots\cdot f_d(t)$ with $k-1$ factors deleted. So $\sigma_*(1)(\Delta_{k-1})\neq 0$ since

$$\Big(\sum_{j=d-k+1}^{d}f_j'(1)\Big)>0 \text{ and } f_1(1)\cdot\cdots\cdot f_{d-k}(1)\neq 0.$$

Therefore, there exists a neighborhood $W\subset U$ of p in M_m on which the radial derivative of Δ_{k-1} is non-zero. W can be chosen to be convex, such that $C_W(m)\subseteq C^R(m)$, and such that on each ray r of M_m which intersects W there exists exactly one conjugate point on $r\cap W$. This follows from the definition of $C^R(m)$ and the property (R3) of F. On each ray r of M_m which intersects W there is then one conjugate point in W necessarily of order k.

Δ_{k-1} is a C^∞ function on W whose differential is everywhere non-zero. Also $C_W(m)=\Delta_{k-1}^{-1}(0)\cap W$. For if $q\in C_W(m)$, then the rank of dF at q is $d-k$ which implies $\Delta_{k-1}(q)=0$. Conversely suppose $\Delta_{k-1}(q)=0$. Let q_1 be the unique conjugate point in W on the ray of M_m through q. Since W is convex the radial line from q_1 to q lies in W. Δ_{k-1} is zero at both q and q_1 and the derivative of Δ_{k-1} along the radial line joining q_1 and q is non zero. Hence $q_1=q$ by the mean value theorem. Therefore $q\in C_W(m)$. Therefore, by the implicit function theorem (1.1), $C_W(m)$ can be given a manifold structure of dimension $d-1$ such that the inclusion $i\colon C_W(m)\to M_m$ is a submanifold in the relative topology. Moreover, since the radial derivative of Δ_{k-1} is non zero, $(M_m)_p=(C^R(m))_p\oplus r_p$ for every $p\in C^R(m)$.

q. e. d.

Definition. Let $p\in C^R(m)$. We define $T(p)$ to be the subspace of the

null space $N(p)$ of $dF(p)$ tangential to the regular conjugate locus at p, i.e., $T(p)=N(p)\cap(C^R(m))_p$.

Consider an open connected submanifold C of the regular conjugate locus $C^R(m)$. The order of the conjugate points comprising C is a constant k. The null space $N(p)$ of $dF(p)$ intersects the tangent space C_p to C in either a k or $k-1$ dimensional subspace which we have denoted by $T(p)$. Let C^k be the subset of points p in C where the dimension of $T(p)$ is k and C^{k-1} the subset where the dimension of $T(p)$ is $k-1$. C^{k-1} is an open subset of C since it is the subset where $d(F\mid C)$ has maximal rank. Moreover $T\mid C^{k-1}$ is an involutive C^∞ distribution and so is $T\mid \mathrm{Int}(C^k)$—indeed the null spaces of the differential of any C^∞ map form an involutive C^∞ distribution on any open subset where they have constant dimension, and $T\mid C^{k-1}$ is the distribution of null spaces of $d(F\mid C^{k-1})$, as is $T\mid \mathrm{Int}(C^k)$ the distribuption of null spaces of $d(F\mid \mathrm{Int}(C^k))$.

THEOREM 3.2. *C^{k-1} is empty for $k\geqq 2$. That is, if p is a regular conjugate point of a regular exponential map $F\colon M_m\to M$ and if the order of p is $\geqq 2$, then the null space $N(p)$ of $dF(p)$ is contained in the tangent space $(C^R(m))_p$ to the regular conjugate locus at p.*

Proof. Suppose C^{k-1} is not empty. We show $k=1$. Let p be a point in C^{k-1}. It would suffice to show that the dimension of the range of $d^2F\mid (M_m)_p\# N(p)$ is $\leqq 1$, $(d^2F\colon (M_m)^2{}_p/(M_m)_p\to M^2{}_{F(p)}/dF((M_m)_p))$. For then, since (R2) implies $d^2F\mid(r_p\# N(p))$ is $1{:}1$, the dimension of $N(p)$ must be $\leqq 1$. Hence $\dim N(p)=1$, whence $k=1$. The dimension of $(M_m)_p\# N(p)$ is $dk-(k(k-1)/2)$. We show the dimension of the kernel of $d^2F\mid((M_m)_p\# N(p))$ is at least $dk-k(k-1)/2-1$ by showing the kernel contains $C_p\# N(p)$. Since $\dim C_p\cap N(p)=k-1$,

$$\dim C_p\# N(p)=(d-1)k-((k-1)(k-2)/2)$$
$$=dk-(k(k-1)/2)-1.$$

Let $X(p)\in C_p$ and $Y(p)\in N(p)$. $X(p)$ can be extended to be a vector field X on a neighborhood U of p in M_m such that whenever $q\in U\cap C$, $X(q)\in C_q$. Similarly $Y(p)$ can be extended to be a vector field Y on U such that whenever $q\in U\cap C$, $Y(q)\in N(q)$. Then $(XY)(p)\in (M_m)^2{}_p$ and $X(p)\# Y(p)=(XY)(p)+(M_m)_p$ as elements of $(M_m)^2{}_p/(M_m)_p$. To show $d^2F(X(p)\# Y(p))=0$ in $M^2{}_{F(p)}/dF((M_m)_p)$, it suffices to show $d^2F((XY)(p))=0$ in $M^2{}_{F(p)}$. But

$$d^2F((XY)(p))(g)=X(Y(g\circ F)(p)=0,$$

where g is any C^∞ function on a neighborhood of $F(p)$, because of the way the extensions X and Y were chosen. $C_p \# N(p)$ is therefore in the kernel of $d^2F \mid ((M_m)_p \# N(p))$ since the generators of $C_p \# N(p)$ are. q. e. d.

In the context of Finsler spaces this last theorem generalizes a result of J. H. C. Whitehead who had proved in [11] that this holds for k greater than half the dimension of the manifold. As Whitehead's proof for the case $k > d/2$ is very geometric in contrast with our proof above, we give a sketch of his proof in Section 4 where we treat Finsler spaces.

So by Theorem 3.2, if C is an open connected submanifold of the regular conjugate locus consisting of conjugate points of order $k \geqq 2$, then T equals N and is a k dimensional involutive C^∞ distribution on C. If $k=1$, then $T \mid \mathrm{Int}(C^1)$ is a one dimensional involutive C^∞ distribution. In both these situations the maximal integral manifolds of T are closed submanifolds of C (resp. $\mathrm{Int}(C^1)$) and have the relative topology. This is easily seen by observing firstly that each of these integral manifolds is mapped to a single point under F and, secondly, that if one chooses a coordinate system locally on C (or $\mathrm{Int}(C^1)$) such that the integral manifolds of T are slices, then F is locally one to one on any slice of the coordinate system complementary to T.

With this last theorem we are now in a position to give normal forms for regular exponential maps on neighborhoods of regular conjugate points of order $\geqq 2$ and for certain of the order 1 cases.

THEOREM 3.3. *Let $F\colon M_m \to M$ be a regular exponential map and $p \in M_m$ a regular conjugate point of F.*

a) *If order $(p) = k \geqq 2$, there exists coordinate systems $x_1, \cdots, x_d$ and $y_1, \cdots, y_d$ on neighborhoods of p and $F(p)$ respectively such that*

$$y_i \circ F = x_i, \qquad i = 1, \cdots, d-k,$$
$$y_i \circ F = x_1 \cdot x_i, \quad i = d-k+1, \cdots, d.$$

b) *If order $(p) = 1$ and if for all points q in a neighborhood of p on the regular conjugate locus $N(q) = T(q)$, i.e. $N(q) \subset (C^R(m))_q$, then there exist coordinate systems $x_1, \cdots, x_d$ and $y_1, \cdots, y_d$ on neighborhoods of p and $F(p)$ respectively such that*

$$y_i \circ F = x_i, \qquad i = 1, \cdots, d-1,$$
$$y_d \circ F = x_1 \cdot x_d.$$

c) *If order $(p) = 1$ and if $T(p) = 0$, i.e. $N(p) \not\subset (C^R(m))_p$, then*

there exist coordinate systems $x_1,\cdots,x_d$ *and* $y_1,\cdots,y_d$ *on neighborhoods of* p *and* $F(p)$ *respectively such that*

$$y_i\circ F=x_i,\qquad i=1,\cdots,d-1,$$
$$y_d\circ F=(x_d)^2.$$

Proof. We prove parts a) and b) together. They were stated separately because the additional hypothesis needed for b), namely that the null space $N(q)$ is tangent to the conjugate locus for all conjugate points q in some neighborhood of p on the regular conjugate locus, holds for regular conjugate points of order $\geqq 2$ by Theorem 3.2.

So suppose p is a regular conjugate point of order k, $1\leqq k\leqq d-1$, and that C is an open connected submanifold of the regular conjugate locus containing p and on which the hypotheses of b) are satisfied in the case $k=1$. There exists a centered coordinate system $u_1,\cdots,u_d$ on a connected neighborhood of p which intersects the conjugate locus in C (or a suitable open subset of C which we still denote by C) such that C is the slice $u_1=0$ and such that the integral manifolds of the involutive distribution $T\mid C=N\mid C$ are given by $u_1=0$ and $u_2=\text{constant},\cdots,u_{d-k}=\text{constant}$. There exists a centered coordinate system $\psi\colon v_1,\cdots,v_d$ on a connected neighborhood of $F(p)$ such that the image under F of the slice

$$u_1=0=u_d=u_{d-1}=\cdots=u_{d-k+1}$$

is the slice

$$v_1=0=v_d=v_{d-1}=\cdots=v_{d-k+1},$$

and such that

$$dF(\partial/\partial u_i(p))=(\partial/\partial v_i)(F(p)),\qquad i=1,\cdots,d-k.$$

Throughout the proof we will neither label coordinate neighborhoods nor be specific about the exact domains and ranges of the maps involved—any change of coordinates or application of Taylor's Theorem may require suitable restrictions of domains or may require that the coordinate neighborhoods be convex and these requirements can be satisfied at each stage of the proof.

With this initial choice of coordinates we have by Taylor's Theorem with integral remainder,

$$v_i\circ F=u_i+R_i,\qquad \operatorname{ord}R_i(p)\geqq 2,\qquad \text{for } i=1,\cdots,d-k,$$
$$\operatorname{ord}(v_i\circ F)(p)\geqq 2,\qquad \text{for } i=d-k+1,\cdots,d.$$

Moreover

$$R_1 \mid (u_1 = 0) \equiv 0 \text{ and } v_i \circ F \mid (u_1 = 0) \equiv 0, \text{ for } i = d - k + 1, \cdots, d.$$

The order (ord) of a function is $j > 0$ at a point if the function vanishes there and is annihilated by all tangent vectors there of order $< j$ but is not annihilated by some tangent vector of order j.

Let

$$w_i = u_i + R_i, \quad \text{for } i = 1, \cdots, d - k,$$
$$w_i = u_i, \qquad \text{for } i = d - k + 1, \cdots, d.$$

Since $\operatorname{ord} R_i(p) \geqq 2$, $\phi: w_1, \cdots, w_d$ is a good centered coordinate system on some neighborhood of p. Since in addition $R_1 \mid (u_1 = 0) \equiv 0$, the coordinate neighborhood can be chosen so that u_1 and w_1 have the same zeros there. We now have

$$v_i \circ F = w_i, \qquad \text{for } i = 1, \cdots, d - k, \tag{3.1}$$

$$\operatorname{ord}(v_i \circ F)(p) \geqq 2, \qquad \text{for } i = d - k + 1, \cdots, d, \tag{3.2}$$

$$v_i \circ F \mid (w_1 = 0) \equiv 0, \qquad \text{for } i = d - k + 1, \cdots, d. \tag{3.3}$$

Let $\pi_{d-k}\colon R^d \to R^d$ be the projection such that

$$\pi_{d-k}(a_1, \cdots, a_d) = (a_1, \cdots, a_{d-k}, 0, \cdots, 0).$$

Define functions y_i on a neighborhood of $F(p)$ by

$$\begin{aligned} y_i &= v_i, && \text{for } i = 1, \cdots, d - k, \\ y_i &= v_i - v_i \circ F \circ \phi^{-1} \circ \pi_{d-k} \circ \psi, && \text{for } i = d - k + 1, \cdots, d. \end{aligned} \tag{3.4}$$

Because of (3.2) this is a good change of coordinates on a neighborhood of $F(p)$. By (3.1) for $i = 1$, and by (3.3) and (3.4) we have

$$y_i \circ F \mid (w_1 = 0) \equiv 0, \qquad \text{for } i = d - k + 1, \cdots, d. \tag{3.5}$$

Therefore there exist functions x_i on a neighborhood of p such that

$$y_i \circ F = w_1 \cdot x_i, \qquad \text{for } i = d - k + 1, \cdots, d.$$

Moreover, $x_i(p) = 0$ since (3.2) and (3.4) imply

$$\operatorname{ord}(y_i \circ F)(p) \geqq 2, \qquad \text{for } i = d - k + 1, \cdots, d. \tag{3.6}$$

Finally, define

$$x_i = w_i, \qquad \text{for } i = 1, \cdots, d - k. \tag{3.7}$$

Then on a neighborhood of p

$$y_i \circ F = x_i, \qquad \text{for } i = 1, \cdots, d-k, \tag{3.8}$$

$$y_i \circ F = x_1 \cdot x_i, \qquad \text{for } i = d-k+1, \cdots, d. \tag{3.9}$$

We must finally show that $x_1, \cdots, x_d$ is a coordinate system on a neighborhood of p. For this we will once again need to use the property (R2) of F. We must show that the

$$\operatorname{rank}\{(\partial x_i/\partial w_j)(p)\} = d, \qquad i, j \in \{1, \cdots, d\}.$$

Due to (3.7), we must show that the

$$\operatorname{rank}\{(\partial x_i/\partial w_j)(p)\} = k, \qquad i, j \in \{d-k+1, \cdots, d\}.$$

Now for $i, j = d-k+1, \cdots, d$, using the fact that our coordinate systems are centered, we obtain from (3.9) and (3.7)

$$(\partial^2(y_i \circ F)/\partial w_1 \partial w_j)(p) = (\partial x_i/\partial w_j)(p).$$

So we must show that

$$\operatorname{rank}\{(\partial^2(y_i \circ F)/\partial w_1 \partial w_j)(p)\} = k, \quad i, j \in \{d-k+1, \cdots, d\}. \tag{3.10}$$

Let $\alpha : M^2{}_{F(p)} \to M^1{}_{F(p)}$ be the natural projection obtained from the coordinate system $y_1, \cdots, y_d$. By (R2) the linear map $\alpha \circ d^2F : (M_m)^2{}_p \to M^1{}_{F(p)}$ must be a surjection. In terms of the coordinate systems $\{y_i\}$ and $\{w_i\}$, $\alpha \circ d^2F$ is represented by the $d \times (d + d(d+1)/2)$ matrix

$$\{(\partial y_i \circ F/\partial w_j)(p) \,|\, (\partial^2 y_i \circ F/\partial w_r \partial w_s)(p)\}, \qquad i, j, r < s \in \{1, \cdots, d\}, \tag{3.11}$$

which, by (R2), has rank d.

Now by (3.7) and (3.8), for $i = 1, \cdots, d-k$, we have

$$(\partial y_i \circ F/\partial w_j)(p) = \delta_{ij}, \qquad j = 1, \cdots, d, \tag{3.12}$$

and

$$(\partial^2 y_i \circ F/\partial w_r \partial w_s)(p) = 0, \qquad r < s \in \{1, \cdots, d\}.$$

By (3.6), for $i = d-k+1, \cdots, d$, we have

$$(\partial y_i \circ F/\partial w_j)(p) = 0, \qquad j = 1, \cdots, d,$$

and by (3.9), (3.7), and the fact $w_1(p) = 0$ we have for $i = d-k+1, \cdots, d$,

$$(\partial^2 y_i \circ F/\partial w_r \partial w_s)(p) = 0, \qquad r < s \in \{2, \cdots, d\}.$$

Finally (3.1) and (3.4) imply that $y_i \circ F$ restricted to the slice $0 = w_{d-k+1} = \cdots = w_d$ is identically zero for $i = d-k+1, \cdots, d$ and hence

$$(\partial^2 y_i \circ F/\partial w_1 \partial w_j)(p) = 0, \qquad j = 1, \cdots, d-k.$$

Therefore since the rank of the matrix (3.11) is d, by (R2), and the rank of the submatrix (3.12) is $d-k$, the rank of the submatrix (3.10) of (3.11) must be k, which finishes the proof of parts a) and b).

The normal form for part c) has been proved by Whitney [12, page 394] for the case $d = 2$, and the proof for arbitrary d is the same, nearly word for word. We need only check that our hypotheses guarantee those of Whitney. Specifically we must show that, in Whitney's terminology, p is a *good* singularity of F and is a fold point. That p is a fold point is simply our hypothesis that $T(p) = 0$. That p is a *good* singularity of F follows from the property (R2). We must show that there is a tangent vector $X \in (M_m)_p$ such that, with respect to a choice of coordinates around p and $F(p)$, $X(\det d^1F) \neq 0$. But this has been shown in the proof of Theorem 3.1; for when $k = 1$, $\Delta_0 = \det d^1F$, and the radial derivative of Δ_0 is non-zero.

q. e. d.

In general, C^∞ maps can be 1:1 on neighborhoods of singular points. Consider, for example, the map of R^2 onto R^2 given by $(x, y) \to (x^3, y)$. This mapping is 1:1 and yet singular along the entire y axis. We conclude this section with a proof that a regular exponential map is never a 1:1 map on any neighborhood of any conjugate point. This has been proved, using different methods, by Morse and Littauer in [5] for analytic Finsler spaces and by Savage in [8] for C^∞ Finsler spaces. We give further remarks on their treatments in Section 4.

THEOREM 3.4. *Let $F: M_m \to M$ be a regular exponential map. Let $p \in C(m)$ be any conjugate point of F. Then F is never* 1:1 *on any neighborhood of p.*

Proof. It suffices to prove the theorem for $p \in C^R(m)$ since, by Theorem 3.1, $C^R(m)$ is dense in $C(m)$. Let C be an open connected submanifold of $C^R(m)$ containing p. If the order of p is $k \geqq 2$, then, by Theorem 3.2, T forms a k dimensional involutive C^∞ distribution on C, and the integral manifold of T through p is mapped to a single point by F. If the order of p is 1, either $p \in \mathrm{Int}(C^1)$, in which case there is a one dimensional integral manifold of T through p which collapses to a point under F, or arbitrarily near p there is a point $q \in C^0$, i.e. $T(q) = 0$, and in this case, by Theorem

3.3 part c), F is not one to one on any neighborhood of q. Hence F is not one to one on any neighborhood of p, for any $p \in C(m)$. q. e. d.

4. The exponential map of a Finsler space. We now show that the exponential maps of Finsler spaces, and in particular of Riemannian manifolds, are regular. For this, we first develop some of the elementary properties of such spaces in a setting convenient for proving regularity. In many respects we follow Sternberg's treatment of calculus of variations problems in [10].

To keep notation simple we will make the following identifications. If M is a manifold and $T(M)$ its tangent bundle, we will often consider M imbedded in $T(M)$ as the zero section. Let M_m be a particular tangent space to M. Because M_m is a finite dimensional linear space, it has a natural manifold structure. In fact, if $\{e_i\}$ is a basis of M_m, then the dual basis $\{r_i\}$ is a coordinate system on M_m. Because of this linear structure there is a natural identification of M_m with $(M_m)_p$ for any $p \in M_m$ where $\sum a_i e_i$ in M_m corresponds to $\sum a_i(\partial/\partial r_i)(p)$ in $(M_m)_p$. These identifications are independent of the basis chosen. By a *constant vector field* on M_m we mean one invariant under these identifications. If v is an element of M_m or of $(M_m)_p$ for any $p \in M_m$, we let the corresponding capital letter V denote the constant vector field on M_m determined by v. Finally, by means of the natural embedding of M_m in $T(M)$, we identify the tangent space $(M_m)_p$ to M_m at p with the vertical tangent space $V(T(M)_p)$ to $T(M)$ at p, that is, with the subspace of $T(M)_p$ tangent to the fibre at p. So elements of $(M_m)_p$ can operate, in this way, on functions on $T(M)$.

We will use m_λ, where λ is a real number, to denote both multiplication by λ on real numbers, $m_\lambda: R \to R: a \to \lambda a$, and multiplication by λ on $T(M)$ where $m_\lambda: T(M) \to T(M)$ is defined by: $m_\lambda(m, v) = (m, \lambda v)$ for all (m, v) in $T(M)$. We will also denote $m_\lambda(p)$ by λp for $p \in T(M)$. The particular use will be clear from the context.

Definition. A Finsler Space is a pair (M, F) consisting of a C^∞ manifold M and a function F on $T(M)$ satisfying:

(F1) F is C^∞ on $T(M) - M$.

(F2) $F \circ m_\lambda = m_\lambda \circ F$ for all real numbers $\lambda \geqq 0$.

(F3) $F \geqq 0$ and equals 0 only on the zero section M of $T(M)$.

(F4) The symmetric bilinear form on $(M_m)_p$, $p \neq 0 \in M_m$, defined by $(v_1, v_2)_{(m,p)} = (\frac{1}{2}) V_1 V_2 (F^2 \mid M_m)(p)$, where V_1 and V_2 are the natural extensions of v_1 and $v_2 \in (M_m)_p$ to constant vector fields on M_m, is positive definite for all $p \neq 0 \in M_m$, and $m \in M$.

That the bilinear form in (F4) is symmetric follows immediately from the fact that the Lie bracket of any two constant vector fields on a linear space is zero. It would have been sufficient to assume non-singularity of this bilinear form for positive definiteness is then a consequence of the other properties of F.

If M is a Riemannian manifold and $\| \ \|$ is the norm associated with the inner product on M, then $(M, \| \ \|)$ is clearly a Finsler space.

We assume, for the remainder of this section, that (M, F) is a fixed d-dimensional Finsler space. To get the exponential map of (M, F) we will need to consider the function F^2 on $T(M)$. We let $G = F^2$. By (F1) and (F2), G is a continuous function on $T(M)$ which is C^∞ on $T(M) - M$ and identically zero on M.

PROPOSITION 4.1. *G is C^1 on $T(M)$ and is C^2 if and only if F is the norm of a Riemannian structure in which case G is C^∞ on $T(M)$.*

Proof. Let $x_1, \cdots, x_d$ be a coordinate system on a neighborhood U in M and let π denote the natural projection of $T(M)$ onto M. Then $x_1 \circ \pi, \cdots, x_d \circ \pi, \bar{x}_1, \cdots, \bar{x}_d$ is the usual coordinate system on $\pi^{-1}(U)$ where $\bar{x}_i(m, v) = v_i$ if $v = \sum v_i (\partial/\partial x_i)(m)$. By (F2), G satisfies

$$(4.1) \qquad G \circ m_\lambda = m_{\lambda^2} \circ G \qquad \text{for } \lambda \geqq 0.$$

One easily computes from (4.1) that the partials of G with respect to the above coordinate system exist and equal 0 at all points $(m, 0)$ in $\pi^{-1}(U)$, and moreover that

$$(4.2) \qquad (\partial G/\partial(x_i \circ \pi)) \circ m_\lambda = m_{\lambda^2} \circ (\partial G/\partial(x_i \circ \pi))$$

and

$$(4.3) \qquad (\partial G/\partial \bar{x}_i) \circ m_\lambda = m_\lambda \circ (\partial G/\partial \bar{x}_i) \text{ on } \pi^{-1}(U) \text{ for } \lambda \geqq 0,$$

which implies the partials are continuous at the zero section. Therefore G is C^1 on $T(M)$.

If F is the norm of a Riemannian structure, G is C^∞. Conversely, suppose G is C^2. We show F is the norm of a Riemannian structure. Let $m \in M$. It follows from (4.1) and the assumption G is C^2 that

$$(4.4) \qquad V_1 V_2 (G \mid M_m) \circ m_\lambda \mid M_m = V_1 V_2 (G \mid M_m) \text{ for } \lambda \geqq 0,$$

for any constant vector fields V_1 and V_2 on M_m. Hence $V_1 V_2 (G \mid M_m)$ is a constant function on M_m everywhere equal to its value at the origin. Therefore at all points in $\pi^{-1}(U)$, the partial derivatives $(\partial^\alpha G/\partial \bar{x}^\alpha)$, for

$|\alpha| \geqq 3$, all vanish. Moreover, formula (4.2) holds for all partial derivatives $\partial^\alpha G/\partial(x\circ\pi)^\alpha$, $|\alpha| \geqq 0$; and so G is C^∞ on $T(M)$. We define a bilinear form $(\ ,\ )$ on M_m by

$$(v_1, v_2) = (\tfrac{1}{2}) V_1 V_2 (G \mid M_m)(0)$$

for v_1 and $v_2 \in M_m$. This bilinear form is symmetric and by (4.4) and (F4), $(v_1, v_2) = (V_1(p), V_2(p))_{(m,p)}$ for any $p \neq 0 \in M_m$. Hence $(\ ,\ )$ is positive definite by (F4) and is an inner product on M_m. This choice of an inner product on each tangent space to M is smooth since G is C^∞ and therefore is a Riemannian metric. Finally, by Taylor's theorem,

$$G(m, v) = (\tfrac{1}{2}) VV(G \mid M_m)(0)$$

for each $(m, v) \in T(M)$, where V is the constant vector field on M_m determined by v. So G is the quadratic form associated with the inner product $(\ ,\ )$ and F is the norm of this Riemannian structure. q. e. d.

There is a natural vertical vector field Z on $T(M)$ defined as follows. If $(m, v) \in T(M)$, then $Z(m, v)$ is the unique element of $(M_m)_v$, that is, of $V(T(M)_{(m,v)})$, which gets identified with v under the natural identification of $(M_m)_v$ with M_m. That is, $Z(m, v) = V(v)$, where V is the constant vector field on M_m associated with v.

One easily computes from (4.1) that

$$ZG = 2G \text{ on } T(M). \tag{4.5}$$

LEMMA 4.2. *Let $v_1 \in (M_m)_v$ where $v \neq 0 \in M_m$. Then*

$$v_1(G) = 2(v_1, Z)_{(m,v)}.$$

Proof. Let V_1 be the constant vector field on M_m associated with v_1. We first show that $[V_1, Z \mid M_m] = V_1$. Let $\{e_i\}$ be a basis of M_m and $\{r_i\}$ the dual basis. Then $V_1 = \sum a_i(\partial/\partial r_i)$ where the a_i are constants, and $Z \mid M_m = \sum r_i(\partial/\partial r_i)$.

$$\begin{aligned}[V_1, Z \mid M_m] &= [\textstyle\sum a_i(\partial/\partial r_i), \sum r_j(\partial/\partial r_j)] \\ &= \sum_{i,j} a_i[(\partial/\partial r_i), r_j(\partial/\partial r_j)] \\ &= \textstyle\sum a_i(\partial/\partial r_i) = V_1.\end{aligned} \tag{4.6}$$

To prove the lemma we simply compute as follows:

$$
\begin{aligned}
2v_1(G) &= 2V_1(G)(m,v) \\
&= V_1(ZG)(m,v) && \text{by (4.5)} \\
&= (Z \mid M_m)(V_1G)(m,v) + [V_1, Z \mid M_m](G)(m,v) \\
&= 2(Z(m,v), V_1(v))_{(m,v)} + V_1(G)(m,v) && \text{by (4.6)} \\
&= 2(Z, v_1)_{(m,v)} + v_1(G).
\end{aligned}
$$

Therefore, $v_1(G) = 2(v_1, Z)_{(m,v)}$. q. e. d.

G induces a map G^* of $T(M)$ into the dual bundle $T^*(M)$. If $(m, v_0) \in T(M)$, $G^*(m, v_0)$ is the element of $M_m{}^*$ defined by:

$$(4.7) \qquad (G^*(m, v_0))(v) = dG_{(m,v_0)}(v) = V(G)(m, v_0)$$

for each $v \in M_m$, where V is the extension of v to a constant vector field on M_m. G^* is a continuous fibre preserving map which on $T(M) - M$ is C^∞. By (4.1) G^* satisfies

$$(4.8) \qquad G^* \circ m_\lambda = m_\lambda \circ G^*, \qquad \text{for } \lambda \geqq 0.$$

We now show G^* is a non singular map on $T(M) - M$. Let $v_1 \in (M_m)_{v_0}$, $v_0 \neq 0 \in M_m$. Then $dG^*_{(m,v_0)}(v_1)$ is tangent to the fibre $M_m{}^*$ at $G^*(m, v_0)$. If we consider an element v of M_m as a linear function on $M_m{}^*$, then $dG^*_{(m,v_0)}(v_1)$ acts on v, and

$$
(4.9) \qquad
\begin{aligned}
(dG^*_{(m,v_0)}(v_1))(v) &= v_1(v \circ G^*) \\
&= v_1(V(G)) && \text{by (4.7)} \\
&= 2(v_1, V(v_0))_{(m,v_0)}.
\end{aligned}
$$

Thus by (4.9) and the non-singularity of the bilinear form $(\ ,\ )_{(m,v_0)}$ due to (F4), if $dG^*_{(m,v_0)}(v_1) = 0$ then $v_1 = 0 \in (M_m)_{v_0}$. Moreover, if $t \in (T(M))_p$ and $d\pi(t) \neq 0$, then $dG^*_p(t) \neq 0$ because $\pi^* \circ G^* = \pi$, where π^* is the natural projection $\pi^*: T^*(M) \to M$. Therefore G^* is non-singular on $T(M) - M$.

There is a canonical non-singular 2-form $d\omega$ on $T^*(M)$ where ω is the 1-form defined by:

$$(4.10) \qquad \omega(t) = v^*(d\pi^*(t)) \text{ for } t \in (T^*(M))_{(m,v^*)}.$$

If $x_1, \cdots, x_d$ is a coordinate system on a neighborhood U in M then $x_1 \circ \pi^*, \cdots, x_d \circ \pi^*, y_1, \cdots, y_d$ is the usual coordinate system on $(\pi^*)^{-1}(U)$, where $y_i(m, v^*) = v_i$ if $v^* = \sum v_i dx_i(m)$. With respect to this coordinate system $d\omega$ has the form

$$(4.11) \qquad d\omega = \sum dy_1 \wedge d(x_i \circ \pi^*).$$

Definition. Let μ be the 2-form $\delta(G^* \mid T(M) - M)(d\omega)$ on $T(M) - M$.

Let $\alpha_{(m,v)}$ be the linear map of $T(M)_{(m,v)}$ onto the vertical subspace $V(T(M)_{(m,v)})$ of $T(M)_{(m,v)}$, which we have identified with $(M_m)_v$, obtained by first projecting $T(M)_{(m,v)}$ onto M_m by $d\pi$ and then following by the natural identification of M_m with $(M_m)_v$.

PROPOSITION 4.3. *μ is a non-singular 2-form on $T(M) - M$ and satisfies:*

(i) $\delta m_\lambda \mu = \lambda\mu$ *for* $\lambda > 0$.

(ii) $\mu(v, t) = (v, \alpha_p(t))_{(p)}$ *for all* $v \in V(T(M)_p)$ *and* $t \in T(M)_p$ *where* $p \in T(M) - M$.

Proof. μ is non-singular since $d\omega$ is and since $G^* \mid T(M) - M$ is a non-singular map. It follows immediately from the definition (4.10) of $d\omega$ that

$$\delta m_\lambda \omega = \lambda\omega \qquad \text{for } \lambda \geqq 0. \tag{4.12}$$

Now assume $\lambda > 0$. Then, with G^* restricted to $T(M) - M$, we have

$$\begin{aligned}
\delta m_\lambda \mu &= \delta m_\lambda (\delta G^* d\omega) = \delta(m_\lambda \circ G^*) d\omega \\
&= d(\delta(m_\lambda \circ G^*)\omega) = d(\delta(G^* \circ m_\lambda)\omega) && \text{by (4.8)} \\
&= d(\delta G^* \delta m_\lambda \omega) = d(\delta G^* \lambda\omega) && \text{by (4.12)} \\
&= \lambda \delta G^* d\omega = \lambda\mu
\end{aligned}$$

which proves (i).

To prove (ii), choose a coordinate system on a neighborhood U of $\pi(p)$ and take the corresponding natural coordinate system on $\pi^{-1}(U)$. Extend v and t to be vector fields V and T on $\pi^{-1}(U)$ by taking the coefficients with respect to these cordinate fields to be constant functions on $\pi^{-1}(U)$. Then $[V, T] = 0$, and V is a vertical vector field on $\pi^{-1}(U)$, i. e., $d\pi(V) = 0$, which when restricted to $M_{\pi(p)}$ is the constant vector field on $M_{\pi(p)}$ associated with $v \in V(T(M)_p)$. T projects to a vector field $d\pi(T)$ on U and the vertical vector field $\alpha(T)$ on $\pi^{-1}(U)$ is a constant vector field on each fibre of $\pi^{-1}(U)$. Since $p \in T(M) - M$, G^* is a diffeomorphism on some neighborhood W of p. So restricting everything to $W \cap \pi^{-1}(U)$ we obtain:

$$\begin{aligned}
\mu(v, t) &= \mu(V, T)(p) = d\omega(dG^*V, dG^*T)(G^*(p)) \\
&= (\tfrac{1}{2})\{dG^*V(\omega(dG^*T)) - dG^*T(\omega(dG^*V)) \\
&\qquad\qquad - \omega[dG^*V, dG^*T]\}(G^*(p)) \\
&= (\tfrac{1}{2}) dG^*V(\omega(dG^*T))(G^*(p)) \\
&= (\tfrac{1}{2}) V(\omega(dG^*T) \circ G^*)(p) \\
&= (\tfrac{1}{2}) V(\alpha(T)(G))(p) && \text{by (4.10) and (4.7)} \\
&= (v, \alpha_p(t))_{(p)} && \text{q. e. d.}
\end{aligned}$$

Definition. Let X be the vector field on $T(M)$ defined to be zero on the zero section M of $T(M)$ and defined on $T(M)-M$ to be the unique vector field such that

(4.13) $$\mu(X,Y) = -(\tfrac{1}{2})Y(G)$$

for all C^∞ vector fields Y on $T(M)-M$.

PROPOSITION 4.4. *X is C^1 on $T(M)$, C^∞ on $T(M)-M$, and is a spray. That is,*

(i) $d\pi(X(m,v)) = v$, *for* $(m,v)\in T(M)$.

In other words, $\alpha(X)=Z$ *on* $T(M)$.

(ii) $\lambda dm_\lambda(X(m,v)) = X(m,\lambda v)$, *for* $(m,v)\in T(M)$ *and* $\lambda \geqq 0$.

Furthermore

(iii) *The Lie derivative* $\theta(X)\mu$ *of* μ *with respect to* X *is* 0.

Proof. That X is C^∞ on $T(M)-M$ is clear from the definition. The statement (i) is trivial for points $(m,0)\in T(M)$. So assume $(m,v)\in T(M)-M$. For any $v_1\in V(T(M)_{(m,v)})$,

$$\mu(v_1,X)(m,v) = (v_1,\alpha_{(m,v)}(X))_{(m,v)}$$

by Proposition 4.3. On the other hand,

$$\begin{aligned}\mu(v_1,X)(m,v) &= (\tfrac{1}{2})v_1(G)(m,v)\\ &= (v_1,Z)_{(m,v)}\end{aligned}$$

by the definition of X and by Lemma 4.2. Therefore

$$(v_1,Z)_{(m,v)} = (v_1,\alpha_{(m,v)}(X))_{(m,v)} \text{ for all } v_1\in (M_m)_v.$$

Since this bilinear form is non singular by (F4), we must have $Z(m,v) = \alpha_{(m,v)}(X)$, which by the natural identification of $(M_m)_v$ with M_m says that $v = d\pi(X(m,v))$ proving (i).

Part (ii) is trivial for $\lambda=0$ or for points $(m,0)\in T(M)$. So suppose $\lambda>0$ and $p\in T(M)-M$. Since μ_p is non-singular, we need only show that $\mu_p(v,\lambda dm_\lambda X) = \mu_p(v,X)$ for all $v\in T(M)_p$. But indeed,

$$\begin{aligned}\mu_p(v,\lambda dm_\lambda X) &= \mu_p(dm_\lambda(dm_{\lambda^{-1}}v),\lambda dm_\lambda X)\\ &= \lambda(\delta m_\lambda \mu)_{p/\lambda}(dm_{\lambda^{-1}}v, X)\\ &= \lambda^2\mu_{p/\lambda}(dm_{\lambda^{-1}}v,X) \qquad \text{by Proposition 4.3},\\ &= (\lambda^2/2)dm_{\lambda^{-1}}(v)(G)(p/\lambda)\\ &= (\lambda^2/2)v(G\circ m_{\lambda^{-1}})\\ &= (\lambda^2/2)v((1/\lambda^2)G) \qquad \text{by (4.1)},\\ &= (\tfrac{1}{2})v(G) = \mu_p(v,X),\end{aligned}$$

which, as remarked above, proves that the vector fields X and $\lambda dm_\lambda X$ are identical on $T(M)$ and therefore proves (ii).

For (iii) we use the well known expression for the Lie derivative in terms of exterior differentiation and interior multiplication. Namely, $\theta(X) = i(X)d + di(X)$ where $i(X)$ is interior multiplication by X. Now $d\mu = 0$ since $\mu = \delta(G^* \mid T(M) - M)(d\omega)$ and $d^2 = 0$. Moreover $d(i(X))\mu = d(-dG) = 0$. Therefore $\theta(X)\mu = 0$.

We finally show that X is C^1 on $T(M)$. Let $x_1, \cdots, x_d$ be a coördinate system on a neighborhood U in M and let $x_1 \circ \pi, \cdots, x_d \circ \pi, \bar{x}_1, \cdots, \bar{x}_d$ be the usual coordinate system on $\pi^{-1}(U)$. Then

$$X = \sum \bar{x}_i(\partial/\partial(x_i \circ \pi)) + \sum b_i(\partial/\partial \bar{x}_i)$$

for some functions b_i on $\pi^{-1}(U)$. The coefficient of X with respect to $(\partial/\partial(x_i \circ \pi))$ is $\bar{x}_i$ by part (i) of this proposition, and $\bar{x}_i$ is C^∞ on $\pi^{-1}(U)$. Because of (ii) and the fact that $dm_\lambda((\partial/\partial\bar{x}_i)(m,v)) = \lambda(\partial/\partial\bar{x}_i)(m,\lambda v)$, for $(m,v) \in \pi^{-1}(U)$, b_i satisfies $b_i \circ m_\lambda = m_{\lambda^2} \circ b_i$ on $\pi^{-1}(U)$. By the same argument as used for G in Proposition 4.1, we see that b_i is C^1 on $\pi^{-1}(U)$ for $i = 1, \cdots, d$. Therefore X is C^1 on $T(M)$. q. e. d.

Assume X is complete and let U_t be the global one parameter group determined by X. U_t, for each real number t is a C^1 diffeomorphism of $T(M)$ and a C^∞ diffeomorphism of $T(M) - M$. It follows from part (ii) of Proposition 4.4 that

$$\pi \circ U_t(m, sv) = \pi \circ U_{st}(m, v) \qquad \text{for } s \geqq 0. \tag{4.14}$$

Definition. The *exponential map* $\exp\colon T(M) \to M$ of the Finsler space (M, F) is defined by:

$$\exp = \pi \circ U_1.$$

Let $m \in M$. The restriction $\exp \mid M_m$ is denoted by $\exp_m$.

If X is not complete, there exists, by (4.14), a neighborhood of the zero section in $T(M)$ on which exp is defined and a neighborhood of 0 in M_m, for each $m \in M$, on which $\exp_m$ is defined. We will assume throughout that X is complete—all statements can be suitably modified in the case X is not complete.

exp (resp. $\exp_m$) is C^∞ on $T(M) - M$ (resp. $M_m - 0$) and C^1 on $T(M)$ (resp. M_m). Let $\rho(t)$, $t \in R^+$, be the ray in M_m through (m, v), $v \neq 0$. Then by (4.14), $\exp_m \circ \rho$ is the projection by π of the integral curve $U_t(m, v)$ of X and by part (i) of Proposition 4.4 $(\exp_m \circ \rho)_*(t) \neq 0$ for any $t \in R^+$. Moreover, by part (i) of Proposition 4.4, and by (4.14) and the fact

$\exp_m$ is C^1, $(d\exp_m) \mid (M_m)_0 \colon (M_m)_0 \to M_m$ is the identity map with our usual identifications. The curves in M obtained as the image under exp of such rays $\rho(t)$ in the tangent spaces are called the *geodesics* of the Finsler space. Notice that since all of our homogeneity assumptions throughout have been for $\lambda \geqq 0$, it is not necessarily true that a geodesic traversed in the negative direction is still a geodesic. So geodesics are oriented curves.

Ambrose, Palais and Singer proved in [2] that there is a one-one correspondence between C^∞ sprays and torsionless affine connections. The spray of a Finsler space is only C^1 and there is no naturally associated torsionless affine connection. In the case (M,F) is a Riemannian manifold, the spray X is C^∞ on $T(M)$ and therefore X is the spray of a torsionless affine connection which is shown to be the Riemannian connection in the next section.

In the usual treatment of Finsler spaces one considers F as defining a problem in the calculus of variations. Every smooth curve σ in M has an F length, $\int_{t_0}^{t_1} F(\sigma(t), \sigma_*(t))\,dt$, and one is interested in the existence and properties of curves with minimal F-length. A necessary condition that a curve be minimal is that the Euler equations be satisfied, and in the case of Finsler spaces the Euler equations have unique solution curves parametrized proportional to F-length with given initial value and derivative. These curves define an exponential map. One easily checks that our geodesics, which are parametrized proportional to F-length, satisfy the Euler equations for F and hence our exponential map is the same as the classical one.

The culmination of the elementary theory of Finsler spaces is the basic theorem that if a curve has minimal G-length it must be a geodesic, if it has minimal F-length it must be a geodesic up to a positive reparametrization, and geodesics do minimize F-length locally before conjugate points. This theorem follows much as in the Riemannian case, from the basic Gauss Lemma—see [1] or [9], for example. Although it is not necessary for the remainder of this paper, we prove here the Gauss Lemma for Finsler spaces. In particular, this gives a proof for the Riemannian case where the Gauss Lemma is usually proved using the structural equations of the Riemannian connection. Our proof, however, involves only the structure we have set up on the tangent bundle $T(M)$; namely, it involves only the function F and the non-singular 2-form μ.

Let a and b be positive real numbers and let β be a C^∞ map of the rectangle $[0,a] \times [0,b]$ into $T(M)$ of the following type. By C^∞ we mean β is to be extendable to a C^∞ map on an open set containing the rectangle

Recall U_t is the one parameter group of the spray X. Let u and v be the canonical coordinate functions on the rectangle with corresponding coordinate vector fields U and V. Then β is to satisfy:

$$\beta(0,v) \in T(M) - M, \quad \text{for } v \in [0,b], \tag{4.15}$$

$$\beta(u,v) = U_u(\beta(0,v)), \tag{4.16}$$

$$d\beta(V(0,0))(G) = 0. \tag{4.17}$$

Let

$$\begin{aligned} \sigma_v &= \pi \circ \beta \mid ([0,a] \times v) \\ \tau_u &= \pi \circ \beta \mid (u \times [0,b]). \end{aligned} \tag{4.18}$$

Then the σ_v are the projections to M of the "longitudinal" curves of β and are geodesics. Hence

$$(\sigma_v(u), (\sigma_v)_*(u)) = \beta(u,v).$$

Gauss Lemma. $((\sigma_0)_*(u), (\tau_u)_*(0))_{(\sigma_0(u),(\sigma_0)_*(u))}$ *is a constant independent of* u. *In particular, if the tangent vectors* $(\sigma_0)_*(u)$ *and* $(\tau_u)_*(0)$ *are orthogonal with respect to* $(\ ,\ )_{(\sigma_0(u),(\sigma_0)_*(u))}$ *for some* u *then they are orthogonal for all* u.

Proof. By (4.17), $\mu(d\beta(V(0,0)), X(\beta(0,0))) = 0$. Since the Lie derivative $\theta(X)\mu$ is 0 by (iii) of Proposition 4.4, and since $dU_u(d\beta(V(0,0))) = d\beta(V(u,0))$ by (4.16), we have $\mu(d\beta(V(u,0)), X(\beta(u,0))) = 0$. Therefore

$$d\beta(V(u,0))(G) = 0, \text{ for } u \in [0,a]. \tag{4.19}$$

Hence the rectangle β is tangent along its base curve $\beta(u,0)$ to a constant "G sphere" in $T(M)$. From the definition of μ in terms of $d\omega$ we obtain, using the fact that $[U,V] = 0$,

$$\delta\beta(\mu)(U,V) = (\tfrac{1}{2})(U(\alpha \circ d\beta(V)(G)) - V(\alpha \circ d\beta(U)(G)))$$

But

$$(\alpha \circ d\beta)(U(u,v)) = \alpha_{\beta(u,v)}(X(\beta(u,v))) = Z(\beta(u,v))$$

by (4.16) and (i) of Proposition 4.4. Therefore, by (4.5) we have

$$\delta\beta(\mu)(U,V) = (\tfrac{1}{2})(U(\alpha \circ d\beta(V)(G))) - V(G \circ \beta). \tag{4.20}$$

The Gauss lemma is now obtained by showing the following derivative is 0:

$$
\begin{aligned}
(d/du)&\left(((\sigma_0)_*(u),(\tau_u)_*(0))_{(\sigma_0(u),(\sigma_0)_*(u))}\right)\\
&=U(u,0)(\alpha\circ d\beta(U),\alpha\circ d\beta(V))_{\beta(u,v)}\\
&=U(u,0)(Z,\alpha\circ d\beta(V))_{\beta(u,v)}\\
&=(\tfrac{1}{2})U(u,0)(\alpha\circ d\beta(V)(G)) \qquad \text{by Lemma 4.2}\\
&=(\delta\beta(\mu)(U,V)+V(G\circ\beta))(u,0) \qquad \text{by (4.20)}\\
&=-(\tfrac{1}{2})d\beta(V(u,0))(G)+d\beta(V(u,0))(G) \quad \text{by (4.13)}\\
&=(\tfrac{1}{2})d\beta(V(u,0))(G)=0 \qquad \text{by (4.19). q.e.d.}
\end{aligned}
$$

THEOREM 4.5. *Let (M,F) be a d-dimensional Finsler space and m a point in M. Assume the spray X of (M,F) is complete. Then the exponential map $\exp_m\colon M_m\to M$ is regular.*

Proof. We have already proved (R1) in the remarks after the definition of $\exp_m$. For the well known continuity property (R3) we refer to Morse [4, Lemma 13.1, page 235]. We now prove (R2). Let p be a non-zero point of M_m. We use freely the natural identification of M_m with $(M_m)_p$ and with $V(T(M)_{(m,p)})$ and the natural identification of $M_{\exp_m(p)}$ with $(M_{\exp_m(p)})_{U_1(m,p)}$ and $V(T(M)_{U_1(m,p)})$. Since $N(p)$ is the null space of $d\exp_m\mid(M_m)_p$, $dU_1(N(p))$ is contained in $V(T(M)_{U_1(m,p)})$. We first prove

$$M_{\exp_m(p)}=d\exp_m((M_m)_p)\oplus dU_1(N(p)). \tag{4.21}$$

For this we need only show that the intersection of these two subspaces of $M_{\exp_m(p)}$ is empty. So let $v\in d\exp_m((M_m)_p)\cap dU_1(N(p))$. There exist w_1 and w_2 in $(M_m)_p$ such that $dU_1(w_1)=v$ and $d\exp_m(w_2)=v$. Now

$$\mu(w_1,w_2)=(w_1,\alpha_{(m,p)}w_2)_{(m,p)}=(w_1,0)_{(m,p)}=0$$

by Proposition 4.3, since w_1 and w_2 are both "vertical" tangent vectors. But, by the invariance of μ,

$$
\begin{aligned}
\mu(w_1,w_2)&=\mu(dU_1w_1,dU_1w_2)\\
&=(dU_1w_1,\alpha_{U_1(m,p)}dU_1w_2)_{U_1(m,p)}\\
&=(v,v)_{U_1(m,p)}.
\end{aligned}
$$

Therefore $(v,v)_{U_1(m,p)}=0$ which implies v equals zero since this bilinear form is positive definite. This proves (4.21). Let $Y\in N(p)$ and extend Y to a vector field on a neighborhood of (m,p) in $T(M)$. Then

$$((YZ)(p)+(M_m)_p)\in r_p\#N(p).$$

To prove (R2) it suffices, in view of (4.21), to prove that

$$d^2\exp_m((YZ)(p))=dU_1(Y(p)). \tag{4.22}$$

Let $\gamma(t)$ be a smooth curve in $M_{\exp_m(p)}$ such that $\gamma(0)=U_1(m,p)$ and $\gamma_*(0)=dU_1(Y(p))$. Let f be a smooth function on a neighborhood of $\exp_m(p)$. Then

$$\begin{aligned} d^2\exp_m((YZ)(p))f &= dU_1Y(dU_1Z(f\circ\pi))(U_1(m,p)) \\ &= dU_1Y(X(f\circ\pi))(U_1(m,p)) \\ &= \lim_{t\to 0}\frac{X(f\circ\pi)(\gamma(t))-X(f\circ\pi)(\gamma(0))}{t} \\ &= (\lim_{t\to 0}\frac{d\pi(X(\gamma(t)))-d\pi(X(\gamma)(0)))}{t})f \\ &= (\lim_{t\to 0}\frac{\gamma(t)-\gamma(0)}{t})f \\ &= \gamma_*(0)f = dU_1(Y(p))f. \end{aligned}$$

q. e. d.

As a corollary of Theorem 4.5, the results of Sections 2 and 3 hold for the exponential maps of Finsler spaces. In particular, we have shown in Theorem 3.2 that the only regular conjugate points of $\exp_m$ at which the null space of $d\exp_m$ can be non-tangent to the conjugate locus are those of order 1. That non-tangency occurs in the order 1 case is well known—for example, the two dimensional Riemannian ellipsoid has as its generic conjugate locus an ellipse and the null spaces of the differential of the exponential map are only tangent to this ellipse at the ends of its major and minor axes since the null space in the Riemannian case is always orthogonal to the rays in the tangent space. As we remarked before, J. H. C. Whitehead proved in [11] that the null spaces of the differential of an exponential map $\exp_m$ of a d-dimensional Finsler space (M,F) are always tangent to the conjugate locus at regular conjugate points of order $k>d/2$. The idea of Whitehead's proof is this: We use the notation of Theorem 3.2. We show that $k\leqq d/2$ if C^{k-1} is not empty. Locally, C^{k-1} is a $d-1$ dimensional slice in the d dimensional tangent space M_m, and the integral manifolds of $T\mid C^{k-1}$ are locally $k-1$ dimensional slices in C^{k-1}. Each of these integral manifolds of $T\mid C^{k-1}$ is mapped to a single point of M under $\exp_m$, and locally $\exp_m$ collapses C^{k-1} to a $d-k$ dimensional submanifold of M. We show each tangent space to this submanifold contains a k dimensional subspace. Then one must have $k\leqq d-k$, that is, $k\leqq d/2$. Consider a particular integral manifold of $T\mid C^{k-1}$ which under $\exp_m$ maps to a point m'. The "spray" of rays in the tangent space M_m from the origin to this integral manifold is k dimensional and each ray in the spray is mapped under $\exp_m$ to a curve from m to m' and these curves at m' are all tangent to the submanifold $\exp_m(C^{k-1})$. This last fact, that these curves

at m' are tangent to the submanifold, is the crucial one and comes from the fact that at points p in C^{k-1}, $(M_m)_p = N(p) + (C^{k-1})_p$. Not only are these curves at m' tangent to the submanifold but at most two of them have the same tangent direction at m' since these curves are geodesics and for a Finsler space the geodesics are uniquely specified by initial value and tangent vector. Therefore the tangents to these curves fill out a k dimensional subset of the tangent space to the submanifold $\exp_m(C^{k-1})$ at m', which implies $k \leqq d - k$.

As a corollary of Theorems 3.4 and 4.5, the exponential map $\exp_m$ of a Finsler space is never 1:1 on any neighborhood of any conjugate point. Morse and Littauer proved this in [5] for analytic Finsler spaces. Briefly, the idea for their proof is this: Assume $p \in M_m$ is a conjugate point and assume $\exp_m$ *is* 1:1 on some neighborhood of p. Then $\exp_m$ is an analytic homeomorphism of some neighborhood of p with a neighborhood of $\exp_m(p)$. Let σ denote the geodesic which p determines from m through $\exp_m(p)$, and assume, for convenience, that p is the first conjugate point to m along σ. Then curves which lie suitably close to σ in M can be lifted to the tangent space M_m. The analyticity of $\exp_m$ implies that the lifted curves are sufficiently smooth to use the Gauss Lemma and a standard argument to show that σ minimizes F-length from m to beyond p relative to nearby curves, which contradicts the fact that a geodesic in no sense minimizes F-length beyond the first conjugate point. For the C^∞ case, Savage, in [8], used Sard's theorem to show the lifted curves are still sufficiently smooth to use the same argument.

5. The Riemannian case. Let M be a complete d-dimensional Riemannian manifold. Let $\mathfrak{X}$ be the C^∞ vector field on the tangent bundle which is the spray of the Riemannian connection (cf. [2]). Let m be a point of M and $\exp_m\colon M_m \to M$ the exponential map of the Riemannian manifold, i.e., the exponential map determined by $\mathfrak{X}$. In the non-complete case $\exp_m$ is only defined on a neighborhood of the origin in M_m and our results apply, suitably restated.

THEOREM 5.1. *Exp_m is a regular exponential map.*

Proof. In view of Theorem 4.5, we need only show that the spray X of M considered as a Finsler space (the function F on $T(M)$ in this case is the norm or length function associated with the Riemannian inner product $\langle\ ,\ \rangle$; that is $F(m,v) = \langle v,v\rangle^{\frac{1}{2}}$) is the same as the Riemannian spray $\mathfrak{X}$. We use the notation of Section 4 for the Finsler space structure on $T(M)$. In the Riemannian case the function G $(G(m,v) = \langle v,v\rangle)$ is C^∞ on $T(m)$ (Proposition 4.1), G^* is a diffeomorphism of $T(M)$ with $T^*(M)$ and μ is a non-singular C^∞ 2-form on $T(M)$. Moreover, the bilinear forms $(\ ,\ )_{(m,p)}$

of (F4) for $p \in M_m$ are invariant under the natural identifications of the tangent spaces to M_m and are all equal under these identifications to the inner product $\langle\ ,\ \rangle$ on M_m. Let $p \in T(M)$. It is sufficient to show that $\mu(X(p) - \mathfrak{X}(p), Y) = 0$ for all vectors Y in $(T(M))_p$; for, in this case, the non-singularity of μ would then imply that $X(p)$ equals $\mathfrak{X}(p)$. So let Y be an element of $(T(M))_p$. We let ∇ denote the covariant differentiation defined by the Riemannian connection, and $\bar{U}_t$ the one parameter group of diffeomorphisms of $T(M)$ associated with the Riemannian spray $\mathfrak{X}$. Let a and b be positive real numbers and let u and v denote the canonical coordinate functions on the rectangle $[0,a] \times [0,b]$ with corresponding coordinate vector fields U and V. We choose a C^∞ map β of this rectangle in $T(M)$ such that the initial corner $\beta(0,0)$ is p, such that $d\beta(V(0,0)) = Y$, and such that $\beta(u,v) = \bar{U}_u(\beta(0,v))$. Therefore $\alpha \circ d\beta(U(u,v)) = Z(\beta(u,v))$ since $d\beta(U(u,v)) = \mathfrak{X}(\beta(u,v))$ and $\mathfrak{X}$ is a spray. We denote the "longitudinal" curves $\pi \circ \beta \mid ([0,a] \times v)$ of the rectangle projected to M by σ_v and the "transverse" curves $\pi \circ \beta \mid (u \times [0,b])$ by τ_u. The σ_v are geodesics of the Riemannian manifold M. We show that the following functions are equal on the rectangle.

$$U(\alpha \circ d\beta(V)(G)) = V(G \circ \beta). \tag{5.1}$$

The left side is equal to $2U((\alpha \circ d\beta(V), Z)_{\beta(u,v)})$ according to Lemma 4.2. In terms of the Riemannian metric this is equal to $2U(\langle \tau_*, \sigma_* \rangle \circ \beta)$, which is equal to $2(\langle \nabla_{\sigma_*}\tau_*, \sigma_* \rangle + \langle \tau_*, \nabla_{\sigma_*}\sigma_* \rangle) \circ \beta$ since the covariant derivative of the metric tensor $\langle\ ,\ \rangle$ is zero. Since $\nabla_{\sigma_*}\sigma_* = 0$ and $\nabla_{\tau_*}\sigma_* = \nabla_{\sigma_*}\tau_*$ because $[U, V] = 0$ and the Riemannian connection has zero torsion, our expression is equal to $2\langle \nabla_{\tau_*}\sigma_*, \sigma_* \rangle \circ \beta$ which equals $V(\langle \sigma_*, \sigma_* \rangle \circ \beta)$. Again using the bilinear forms of the Finsler space structure, this becomes $V((Z,Z)_{\beta(u,v)})$ which, according to Lemma 4.2, equals $V((\frac{1}{2})Z(G) \circ \beta)$. Finally, by (4.5), this equals the right hand side of (5.1). We now show $\mu(X(p) - \mathfrak{X}(p), Y) = 0$, or, in other words, that $\mu(X(p), Y) = \mu(\mathfrak{X}(p), Y)$. The left hand side is equal to $(-\frac{1}{2})(Y(G))$ by the definition of X. On the right hand side we use the definition of μ in terms of $d\omega$ on $T^*(M)$ and the fact that $[U, V] = 0$, obtaining

$$\mu(\mathfrak{X}(p), Y) = (\tfrac{1}{2})\{U(\alpha \circ d\beta(V)(G)) - V(\alpha \circ d\beta(U)(G))\}(0,0).$$

The first term of the right hand side of this expression is $(\frac{1}{2})V(0,0)(G \circ \beta)$ by (5.1), and this equals $\frac{1}{2}(Y(G))$. The second term is

$$-(\tfrac{1}{2})V(0,0)(Z(G) \circ \beta)$$

which equals $-Y(G)$ by (4.5). Thus $\mu(X(p) - \mathfrak{X}(p), Y) = 0$. q.e.d.

One can, of course, prove Theorem 5.1 directly from standard facts in

Riemannian geometry, and we indicate how that proof goes. Property (R1) holds since $\exp_m$ is C^∞ in the Riemannian case and $\exp_m$ sends the rays in M_m into the non-trivial geodesics through m the tangent vectors of which never vanish. Property (R3) is the continuity property for which we refer to Morse [4, Lemma 13.1, page 235]. Let γ be a non-trivial geodesic such that $\gamma(0)=m$. Let J_γ be the d-dimensional space of Jacobi fields along γ which vanish at m. (For the definition and properties of Jacobi fields refer to Ambrose [1] or Singer [9]). Let n be a point on γ. If Y is a vector field along γ we use Y' to denote the covariant derivative of Y with respect to γ_*. We define

$$A_n=\{v\in M_n \mid v=Y'(n) \text{ for some } Y\in J_\gamma \text{ such that } Y(n)=0\},$$
$$B_n=\{v\in M_n \mid v=Y(n) \text{ for some } Y\in J_\gamma\}. \tag{5.2}$$

So A_n is the subspace of M_n formed by the covariant derivatives at n of the Jacobi fields along γ vanishing both at m and at n. And B_n is the subspace of M_n formed by the values at n of all Jacobi fields along γ vanishing at m. It is well known that A_n and B_n are orthogonal complements in M_n. Roughly said, Jacobi fields and their derivatives span the tangent space at each point along a geodesic. This fact is the crux of the property (R2) which we prove as follows. Let p be a non zero point in M_m. If p is not a conjugate point, $d\exp_m((M_m)_p)=M_{\exp_m(p)}$, so (R2) holds trivially. Suppose p is a conjugate point, say of order k. Let $A_1,\cdots,A_k$ be k elements of $(M_m)_p$ which span the null space $N(p)$ of $d\exp_m \mid (M_m)_p$, and let $A_1,\cdots,A_k$ denote the corresponding constant vector fields on M_m. Let $\sigma(t)$ be the ray in M_m through p, let γ be the corresponding geodesic $\exp_m\circ\sigma$, and let $n=\exp_m(p)$. Let $Y_i(t)=d\exp_m tA_i(\gamma(t))$. Then the Y_i are k linearly independent Jacobi fields along γ vanishing at m and n. Therefore, the covariant derivatives $Y_i'(1)$ form a basis of A_n. The k elements $\sigma_*(1)\#A_i(p)$ form a basis of $r_p\#N(p)$ and according to Lemma 2.1,

$$d^2\exp_m(\sigma_*(1)\#A_i(p))=Y_i'(1)+d\exp_m((M_m)_p).$$

Since $B_n=d\exp_m((M_m)_p)$ and $M_n=A_n\oplus B_n$, $d^2\exp_m$ maps $r_p\#N(p)$ isomorphically onto $M_n/d\exp_m((M_m)_p)$, which proves (R2) and hence the regularity of $\exp_m$.

As a corollary of Theorem 5.1, the results of Sections 2 and 3 for regular exponential maps hold for the exponential maps of Riemannian manifolds.

Remark. In the Riemannian case, according to the Gauss Lemma, the

null spaces of the differential of the exponential map are always orthogonal to the rays in the tangent spaces. Therefore the integral manifolds of the involutive distributions T of Theorem 3.2 lie on spheres in the tangent space.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
UNIVERSITY OF CALIFORNIA AT BERKELEY.

REFERENCES.

[1] W. Ambrose, "The Cartan structural equations in classical Riemannian geometry," *Journal of the Indian Mathematical Society*, vol. 24 (1960), pp. 23-76.

[2] W. Ambrose, R. S. Palais and I. M. Singer, "Sprays," *Anais da Academia Brasileira de Ciencias*, vol. 32 (1960), pp. 163-178.

[3] C. Chevalley, *Theory of Lie Groups*, Princeton Mathematical Series 8, Princeton University Press, Princeton, N. J., 1946.

[4] M. Morse, *The Calculus of Variations in the Large*, American Mathematical Society Publications, vol. 18, 1934.

[5] M. Morse and S. B. Littauer, "A characterization of fields in the calculus of variations," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 18 (1932), pp. 724-730.

[6] S. B. Myers, "Connections between differential geometry and topology I. Simply connected surfaces," *Duke Mathematical Journal*, vol. 1 (1935), pp. 376-391.

[7] ———, "Connections between differential geometry and topology II. Closed surfaces," *Duke Mathematical Journal*, vol. 2 (1936), pp. 95-102.

[8] L. J. Savage, "On the crossings of extremals at focal points," *Bulletin of the American Mathematical Society*, vol. 49 (1943), pp. 467-469.

[9] I. M. Singer, *Lectures on Differental Geometry*, notes by E. M. Brown, M. I. T., (1962).

[10] S. Sternberg, *Lectures on Differential Geometry*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1964.

[11] J. H. C. Whitehead, "On the covering of a complete space by the geodesics through a point," *Annals of Mathematics* (2), vol. 36 (1935), pp. 679-704.

[12] H. Whitney, "On singularities of mappings of Euclidean spaces. I. Mappings of the plane into the plane," *Annals of Mathematics* (2), vol. 62 (1955), pp. 374-410.

DIOPHANTINE PROBLEMS OVER LOCAL FIELDS I.*

By JAMES AX and SIMON KOCHEN.

0. Introduction. A conjecture of Artin states that every form f of degree d in $n > d^2$ variables over Q_p, the p-adic completion of the rationals, has a non-trivial zero in Q_p. For the case $d = 2$, this is a classical theorem about quadratic forms. A proof of the conjecture for $d = 3$ was given by Lewis in [13].

In this paper we prove:

(1) *For every positive integer d there exists a finite set of primes $A = A(d)$ such that for every prime $p \notin A$ every form f of degree d in $n > d^2$ variables over Q_p has a non-trivial zero in Q_p.*[1]

This and the analogous assertion for the completion of a number field k (here A depends only on d and $[k: Q]$) follow from Theorem 5. A further result obtained is the following:

(2) *Let f be a polynomial without constant term of degree d in $n > d$ variables over the ring Z of rational integers. Then there exists a finite set B of primes such that for every prime $p \notin B$, f has a non-trivial zero in Q_p.*

(See the Corollary to Theorem 1.) This re-proves[2] a conjecture of Lang [12].

These and similar results are special cases of the following general

* While working on this paper the first author was partly supported by the U. S. Army Research Office (Durham) contract number DA-31-124-ARO-D-107 and National Science Foundation Grant GP-2243; the second author by National Science Foundation Grant GP-124. The authors thank W. Feit and D. Lewis for their many useful suggestions.

Received August 24, 1964.

Revised April 15, 1965.

[1] Recently D. J. Lewis and B. J. Birch have proved the special cases $d = 5, 7, 11$ of (1).

[2] After completing the original manuscript, we were informed by D. Lewis that N. Greenleaf has obtained the Corollary to Theorem 1, using algebraic-geometric techniques.

principle. Let S_p be the field of power series (or, more precisely, formal meromorphic series expansions about 0) over the field R_p of p elements Then:

For every elementary statement Δ about valued fields there exists a finite set C of primes such that for every prime $p \notin C$, Δ is valid in Q_p if and only if Δ is valid in S_p.

A precise definition of "elementary statement" will be given in Section 3; however, the rough meaning of the principle is intuitively clear and reflects a well-known feeling among mathematicians that the Q_p are "like" the S_p.

Our methods are a composite of some ideas in logic and (general) valuation theory. From logic we make use of an algebraic construction called ultraproducts. In valuation theory, we introduce the notion of an L-field. Lemma 17 gives a useful sufficient condition for a field to be an L-field. Proposition 1 demonstrates the crucial approximation property of these fields which we need. We combine the notion of L-fields with the theory of pseudo-Cauchy sequences developed in Kaplansky [8]. From the fields Q_p and S_p we construct the ultraproduct fields $\mathfrak{Q}$ and $\mathfrak{S}$. It follows from known results in logic that the above-mentioned meta-mathematical principle is equivalent to the isomorphism of $\mathfrak{Q}$ and $\mathfrak{S}$. The major portion of the paper is devoted to proving the isomorphism of $\mathfrak{Q}$ and $\mathfrak{S}$.

The paper [3] includes alternate proofs of the main results given here and in [2]. These alternate proofs avoid the analytic method involving pseudo-Cauchy sequences.

For the convenience of the reader we have arranged the proofs of the mathematical applications of the above-mentioned principle so that no knowledge of logic is required. Nevertheless, the reader who is (or is willing to become) familiar with the rudiments of logic can effect certain simplifications in the proof. In particular, each of the lemmas in Section 1 which prove that a certain property is preserved under passage to the ultraproduct may be subsumed under the statement (**II**) preceding Theorem 6 by showing that the property is equivalent to a set of elementary statements.

We remark that although the above statement (1) requires Theorem 4, from which the general principle follows, statement (2) is a consequence of very simple properties of ultraproducts, and is proved early in the paper.

1. Ultraproducts of valued fields. We first recall some facts about filters. A family D of non-empty subsets of a set I is called a *filter* on I if

(i) $s, t \in D$ implies $s \cap t \in D$

(ii) $s \in D$ and $s \subseteq t \subseteq I$ imply $t \in D$.

A filter is called *principal* if there is an $r \in D$ such that $D = \{t \mid r \subseteq t \subseteq I\}$. An *ultrafilter* on I is a filter which is maximal among the class of all filters on I. A necessary and sufficient condition that a filter D be an ultrafilter is that, for every $r \subseteq I$, $r \notin D$ implies that $(I - r) \in D$. An ultrafilter is principal if and only if $D = \{t \mid i \in t \subseteq I\}$ for some $i \in I$. A use of Zorn's Lemma shows that every infinite set $s \subseteq I$ lies in some non-principal ultrafilter on I. If $I - s$ is finite (i.e. s is *cofinite*) then s is an element of every non-principal ultrafilter on I.

There is a natural one-one correspondence between the family of all ultrafilters on I and the family of all finitely-additive measures on I which take on the values 0 and 1. The measure μ_D corresponding to the ultrafilter D is defined by the condition that $\mu_D(s) = 1$ if and only if $s \in D$, for all $s \subseteq I$. The measure-theoretic language will be employed here merely as a device for simplifying the statements of results.

Let F be a valued field. We denote by *ord* the valuation function on F. We let $\mathfrak{O}_F = \{c \mid \mathrm{ord}(c) \geqq 0\}$ denote the valuation ring, $\bar{F}$ the residue class field of F, and $\bar{q}$ the residue class in $\bar{F}$ of an arbitrary element q of $\mathfrak{O}_F$. If $f \in F[X]$, we let $\bar{f} \in \bar{F}[X]$ denote the polynomial with coefficients residue classes of the corresponding coefficients of f. A map σ of a valued field F onto a valued field F' is an *analytic isomorphism* if σ is a field isomorphism which induces an isomorphism of $\mathrm{ord}(F)$ onto $\mathrm{ord}(F')$. In addition, if $\mathrm{ord}(F) = \mathrm{ord}(F')$ then we require that the induced map be the identity.

Now let $\{F_i \mid i \in I\}$ be an indexed family of valued fields with valuation groups $\{H_i \mid i \in I\}$. Given an ultrafilter D on I we define a new valued field, called an *ultraproduct* of the valued fields F_i. The field, denoted by $\prod_{i \in I} F_i/D$, and its valuation group, denoted by $\prod_{i \in I} H_i/D$, are each homomorphic images of the corresponding (complete) direct products $\prod_{i \in I} F_i$ and $\prod_{i \in I} H_i$. The elements of $\prod_{i \in I} F_i/D$ are equivalence classes of functions $f \in \prod_{i \in I} F_i$, where $f \equiv g$ if $\{i \mid f(i) = g(i)\} \in D$. We may express this in more suggestive measure-theoretic language by using the finitely-additive measure μ_D defined by D. We may then write:

$$f \equiv g \text{ if and only if } f = g \text{ [a.e.] (with respect to } \mu_D).$$

Similarly, if $\gamma, \delta \in \prod_{i \in I} H_i$, we define

$$\gamma \equiv \delta \text{ if } \gamma = \delta \text{ [a.e.]},$$

and if $f^* \in \prod_{i \in I} F_i/D$ and $\gamma^*, \delta^* \in \prod_{i \in I} H_i/D$ we write

$$\gamma^* < \delta^* \text{ if } \gamma < \delta \text{ [a.e.]}$$

and $\operatorname{ord}(f^*) = \gamma^*$ if $\operatorname{ord}(f) = \gamma$ [a.e.].

That the resulting system is a valued field may be checked either directly or by applying the statement (III) preceding Theorem 6. If D is a principal ultrafilter, say $D = \{t \mid j \in t \subseteq I\}$, for some $j \in I$, then $\prod_{i \in I} F_i/D$ is analytically isomorphic to F_j. Hence, our main interest will lie in the case where D is non-principal; we refer to this case by the term *non-principal* ultraproduct. It is worth remarking that for the case of fields ultraproducts may be defined directly in terms of ideals. For there is a one-one correspondence between the class of all filters on I and the class of all ideals in the ring $\prod_{i \in I} F_i$; furthermore, ultrafilters correspond to maximal ideals and principal filters to principal ideals. (See Kochen [9, Th. 8.1].).

Note. We shall systematically let q^* denote an arbitrary element of $\prod_{i \in I} F_i/D$, where q^* is the homomorphic image of $q \in \prod_{i \in I} F_i$. A similar remark applies to our use of $\gamma^* \in \prod_{i \in I} H_i/D$.

LEMMA 1. *Let $\{F_i \mid i \in I\}$ be a family of valued fields, and D an ultrafilter on I. Then $\overline{\prod_{i \in I} F_i/D} \approx \prod_{i \in I} \bar{F}_i/D$.*

Proof. The isomorphism map ν is given by $\nu(\overline{q^*}) = (\bar{q})^*$, where $\bar{q}$ is defined by $\bar{q}(i) = \overline{q(i)}$. Since the condition $\operatorname{ord}(q^*) \geqq 0$ implies that $\operatorname{ord}(q) \geqq 0$ [a.e.], i.e. $q(i) \in \mathfrak{O}_{F_i}$ [a.e.], we have that $\bar{q}^* \in \prod_{i \in I} \bar{F}_i/D$. By the canonical nature of the map, it is sufficient to show that ν is well-defined and one-one. But $\overline{q^*} = \overline{t^*}$ is equivalent to $\operatorname{ord}(q^* - t^*) > 0$, i.e. $\operatorname{ord}(q - t) > 0$ [a.e.], which is in turn equivalent to $\bar{q} = \bar{t}$ [a.e.], or $(\bar{q})^* = (\bar{t})^*$, as required.

It will be convenient to refer to the following statement about a valued field F as the *Hensel-Rychlik property*:

Let $J \in \mathfrak{O}_F[X]$ be monic, and denote the discriminant of J by $D(J)$. If there exists an $f \in \mathfrak{O}_F$ such that $\operatorname{ord}(J(f)) > \operatorname{ord}(D(J))$, then there is an $f' \in \mathfrak{O}_F$ such that $J(f') = 0$.

LEMMA 2. *If $\{F_i \mid i \in I\}$ is a family of fields which satisfy the Hensel-*

Rychlik property, then every ultraproduct $F^* = \prod_{i\in I} F_i/D$ *also satisfies the Hensel-Rychlik property.*

Proof. Let $\sum_{j=0}^{n} a_j^* X^j \in \mathfrak{O}_{F^*}[X]$ with $a_n^* = 1$. Suppose $f^* \in \mathfrak{O}_{F^*}$ is such that $\operatorname{ord}(\sum_{j=0}^{n} a_j^* f^{*j}) > \operatorname{ord}(D(\sum_{j=0}^{n} a_j^* X^j))$. By our convention, we have that $a_0, \cdots, a_n$, $f \in \prod_{i\in I} F_i$ such that, for $0 \leqq j \leqq n$, $\operatorname{ord}(a_j) \geqq 0$ [a. e.], $a_n = 1$ [a. e.], $\operatorname{ord}(f) \geqq 0$ [a. e.], and $\operatorname{ord}(\sum_{j=0}^{n} a_j f^j) > \operatorname{ord}(D(\sum_{j=0}^{n} a_j X^j))$ [a. e.]. Hence, since filters are closed under finite intersections,

$$\operatorname{ord}(a_j) \geqq 0,\ 0 \leqq j \leqq n,\ a_n = 1,\ \operatorname{ord}(f) \geqq 0,$$
$$\operatorname{ord}(\sum_{j=0}^{n} a_j f^j) > \operatorname{ord}(D(\sum_{j=0}^{n} a_j X^j)) \text{ [a. e.]}.$$

Now, by the hypothesis of the lemma, there exists an $f' \in \prod_{i\in I} F_i$ such that

$$\operatorname{ord}(f') \geqq 0,\ \sum_{j=0}^{n} a_j (f')^j = 0 \quad \text{[a. e.]}.$$

It follows that $\operatorname{ord}(f'^*) \geqq 0$, i. e. $f'^* \in \mathfrak{O}_{F^*}$, and $\sum_{j=0}^{n} a_j^* (f'^*)^j = 0$, as required.

In this paper we shall only use the Hensel-Rychlik property for the more familiar case where $\operatorname{ord}(D(J)) = 0$. We have defined the stronger Hensel-Rychlik property for its later use in [2].

Remark. We recall that the *divisible hull* of a torsion-free abelian group H is $H \otimes_Z Q$, the unique (up to H-isomorphism) torsion-free divisible abelian group $\tilde{H}$ containing H such that $\tilde{H}/H$ is a torsion group. If E is a valued field, we have for any extension of ord to an algebraic closure $\tilde{E}$ of E that

$$\operatorname{ord}(\tilde{E}) = \widetilde{\operatorname{ord}(E)}.$$

We require the following result from valuation theory.

LEMMA 3. *Let F be a valued field which satisfies the Hensel-Rychlik property, and suppose* $\operatorname{char} \bar{F} = 0$. *Let $\psi\colon \mathfrak{O}_F \to \bar{F}$ be the residue class map. If T is a maximal subfield of F on which the valuation* ord *is trivial, then $\psi \mid T$ is an isomorphism of T onto $\bar{F}$.*

Proof. That there exists a maximal subfield T of F on which ord is trivial follows immediately from Zorn's Lemma. It follows from the preceding Remark, that T is relatively algebraically closed in F.

Next we show that the field $\psi(T)$ is relatively algebraically closed

in $\bar{F}$. If not, let $\beta \in \bar{F} - \psi(T)$ be algebraic over $\psi(T)$, and let $g(X) = \sum_{i=0}^{n} a_i X^i \in \psi(T)[X]$ be an irreducible polynomial such that $g(\beta) = 0$. Let $s_i \in \mathfrak{O}_F$ be such that $\psi(s_i) = a_i$. From the Hensel-Rychlik property of F it follows that there is a root of $\sum_{i=0}^{n} s_i X^i$ in F, and hence in T, since T is relatively algebraically closed in F. Hence, $g(X)$ has a root in $\psi(T)$, contradicting the fact that $g(X)$ is irreducible over $\psi(T)$.

We now show that $\psi(T) = \bar{F}$. Otherwise, let $c \in \bar{F} - \psi(T)$. We know from the preceding paragraph that c is transcendental over $\psi(T)$. Take $v \in \mathfrak{O}_F$ such that $\psi(v) = c$. Then $v \notin T$. Let $J \in T[X]$ be monic. If $\operatorname{ord}(J(v)) \neq 0$, then $\operatorname{ord}(J(v)) > 0$, i.e. $\bar{J}(c) = 0$, a contradiction to the transcendentality of c over $\psi(T)$. Thus ord is trivial on $T(v)$. This contradiction to the maximality of T, completes the proof of the lemma.

For p in the set P of rational primes, let R_p denote the field with p elements.

LEMMA 4. *Every non-principal ultraproduct $\mathfrak{R} = \prod_{p \in P} R_p/D$ is a field of characteristic* 0.

Proof. The identity of $\mathfrak{R}$ is the element 1^*, where $1(p) = 1$, the identity element of R_p. If n is a positive integer, then $n \cdot 1^* = n^* \neq 0$ in $\mathfrak{R}$ since $n \neq 0$ in R_p for all but a finite number of p, and hence [a.e.].

As an immediate application of the preceding theorems about ultraproducts, we now prove a result on the refinability of solutions of diophantine equations. For each rational prime p, let Q_p denote the p-adic completion of the field Q of rationals. Let Z_p denote the valuation ring of Q_p, and let Z denote the ring of rational integers. If $\alpha \in R_p$ and $a \in Z_p$ is such that $\bar{a} = \alpha$, we refer to the element a as a *refinement* of α.

THEOREM 1. *Let $f_1, \cdots, f_r$ be polynomials in n variables over Z. Then for all but a finite number of primes p, every solution of $\bar{f}_1 = \cdots \bar{f}_r = 0$ in R_p is refinable to a solution of $f_1 = \cdots = f_r = 0$ in Z_p.*

Proof. Let $\mathfrak{Q}$ denote a non-principal ultraproduct $\prod_{p \in P} Q_p/D$. Since, as is well known, (Rychlik [15]) the Q_p's satisfy the Hensel-Rychlik property, $\mathfrak{Q}$ also has this property by Lemma 2. By Lemma 1 we may identify $\bar{\mathfrak{Q}}$ with the field $\mathfrak{R} = \prod_{p \in P} R_p/D$. Hence, it follows from Lemmas 3 and 4 that there is a monomorphism $\phi\colon \mathfrak{R} \to \mathfrak{Q}$ such that $\overline{\phi(a^*)} = a^*$, for every $a^* \in \mathfrak{R}$.

Now let B be the set of primes p such that there exists a solution $(\alpha_1(p),\cdots,\alpha_n(p))$ of $\bar{f}_1=\cdots=\bar{f}_r=0$ in R_p which is not refinable to a zero in Q_p. We now suppose, contrary to the theorem, that B is infinite. There exists a non-principal ultrafilter D on I with $B\in D$. Define $\alpha_i(p)=0$ for $i=1,\cdots,n$ for all $p\notin B$. Hence $(\alpha_1^*,\cdots,\alpha_n^*)$ is a solution of the system in the corresponding ultraproduct $\mathfrak{R}$. If we let $a_i^*=\phi(\alpha_i^*)$, $1\leqq i\leqq n$, then it follows that $(a_1^*,\cdots,a_n^*)$ is a solution of $f_1=\cdots=f_r=0$ in $\mathfrak{Q}$, and $\overline{a_i^*}=\alpha_i^*$. We conclude that if S is the set of primes p such that $(a_1(p),\cdots,a_n(p))$ is a solution of $f_1=\cdots=f_r=0$ in Z_p and $\overline{a_i(p)}=\alpha_i(p)$, then $S\in D$. If we now take $p\in S\cap B\in D$, then $(a_1(p),\cdots,a_n(p))$ is a zero of $(f_1,\cdots,f_r)$ which is a refinement of the zero $(\alpha_1(p),\cdots,\alpha_n(p))$ in R_p. This contradiction proves the theorem.

COROLLARY.[2] *Let $f_1,\cdots,f_p$ be polynomials without constant term in n variables over Z. Let d_i be the degree of f_i for $1\leqq i\leqq r$. If $n>\sum_{i=1}^{r}d_i$, then there exists a finite set B of primes such that for all $p\notin B$ the f_i have a common zero in Q_p.*

Proof. If in the above corollary Q_p is replaced by R_p, then the resulting assertion (with B empty) is a theorem of Chevalley [4]. (A quick proof of Chevalley's theorem may be found in Ax [1].) The Corollary is now an immediate consequence of Theorem 1.

This Corollary implies the following result conjectured by Lang [12, Sect. 2, p. 243]: A homogeneous polynomial f of degree d in $n>d$ variables with coefficients in the integers $\mathfrak{O}$ of an algebraic number field, has a zero in $\mathfrak{O}_\mathfrak{p}$, for all primes $\mathfrak{p}$ of $\mathfrak{O}$ not in a finite set depending only upon f (and $\mathfrak{O}$).

Definition. We shall say that a field F has the *solvability property* if every finite normal extension of F is solvable.

LEMMA 5. *If $\{F_i\mid i\in I\}$ is a family of fields with the solvability property then every ultraproduct of the F_i has the solvability property.*

Proof. We shall give a proof in accordance with the property (**II**) of ultraproducts mentioned immediately before Theorem 6 by reducing the solvability property of a field to a set of elementary statements. We note, however, that for the elementary statements to which the solvability property is reduced the property (**II**) may be directly verified. Let $Y;U_1,\cdots,U_n;$ $\alpha_1,\cdots,\alpha_n$ be variables, and let

$$\psi(Y;U;\alpha)=\prod(Y-\sum_{i=1}^{n}S(U_i)\alpha_i)\in Z[Y;U;\alpha],$$

where the product is extended over all permutations S of the U_i. Then there exists a unique polynomial $\phi(Y;U;A_0,\cdots,A_{n-1}) \in Z[Y;U;A]$ such that if t_i is the $(n-i)$-th elementary symmetric function of the α_j, then

$$\phi(Y;U;t) = \psi(Y;U;\alpha)$$

Let $K_1,\cdots,K_g$ be all the solvable groups of permutations of the U_i. For any field F the following two statements are equivalent (see Van der Waerden [17, Sec. 61]):

(*) Every splitting field of an equation of degree n is solvable over F.

(**) For all $a_0,\cdots,a_{n-1} \in F$, if $\lambda(Y;U)$ is an irreducible factor of $\phi(Y;U;a_0,\cdots,a_{n-1}) \in F[Y;U]$, then the group of permutations of the U_i fixing $\lambda(Y;U)$ is one the groups K_j, $1 \leqq j \leqq g$.

Thus, the condition (**) holds for each field F_i. This implies that (**) holds in $\prod_{i \in I} F_i/D$, and hence that (*) holds in $\prod_{i \in I} F_i/D$. This proves Lemma 5.

Definition. A *Z-group* is an ordered (additive) abelian group G with a smallest positive element 1 and such that $(G: nG) = n$, for all positive integers n.

It is easily seen that the Euclidean algorithm is valid for any Z-group G: For all $\gamma \in G$ and for all positive integers n there exist a unique $\delta \in G$ and $0 \leqq r < n$ such that $\gamma = n\delta + r$. We indicate the cardinality of a set G by $|G|$.

LEMMA 6. *Let $\{G_i \mid i \in I\}$ be a family of Z-groups.*

(a) *Every ultraproduct $\prod_{i \in I} G_i/D$ is a Z-group.*

(b) *If $|G_i| \leqq 2^{\aleph_0}$, for all $i \in I$, $|I| = \aleph_0$, and D is non-principal, then $|\prod_{i \in I} G_i/D| = 2^{\aleph_0}$.*

Proof. The proof of part (a) is an immediate consequence of the definition.

We turn to condition (b). Assume that I is the set of positive integers and set $I_i = \{1,2,\cdots,i\}$. Let $F_i\colon 2^{I_i} \to G_i$ be any one-one map of the power set 2^{I_i} of I_i into G_i. Now define the map $\lambda\colon 2^I \to \prod_{i \in I} G_i/D$ by setting $\lambda(s) = g_s{}^*$, for $s \subseteq I$, where g_s is defined by $g_s(i) = F_i(s \cap I_i)$. It suffices to prove that λ is a one-one map. Suppose then that s and t are distinct subsets of I. It follows that for all but a finite number of $i \in I$, $s \cap I_i \neq t \cap I_i$, i.e. $g_s(i) \neq g_t(i)$. Since the complement of any finite set lies in the non-principal ultrafilter D, $g_s \neq g_t$ [a.e.], i.e. $g_s{}^* \neq g_t{}^*$.

A slight modification of the above proof (See Kochen [9, Th. 6.5)] shows that $|\prod_{p \in P} R_p/D| = 2^{\aleph_0}$ for D non-principal.

Remark. In [3], we show under the hypothesis of Lemma 6(b) that the group $\prod_{i \in I} G_i/D \approx \prod_{p} (Z_p \oplus Q)$, where the last product is extended over all rational primes p.

Definition. Let F be a valued field with valuation group G. A *cross-section* of F is a monomorphism $\pi\colon G \to F$ such that $\operatorname{ord}(\pi(\gamma)) = \gamma$, for all $\gamma \in G$.

LEMMA 7. *Let $\{F_i \mid i \in I\}$ be a family of valued fields such that there exists a cross-section π_i of each F_i. Then there exists a cross-section of each ultraproduct $\prod_{i \in I} F_i/D$.*

Proof. Define $\pi\colon \prod_{i \in I} G_i/D \to \prod_{i \in I} F_i/D$ by $\pi(\gamma^*) = f^*$, $f(i) = \pi_i(\gamma(i))$ for $i \in I$. Then it is easily checked that π is a cross-section.

Definition. A valued field F has the *uniqueness property* if the valuation ord has a *unique* extension to an algebraic closure $\tilde{F}$ of F.

LEMMA 8. *If $\{F_i \mid i \in I\}$ is a family of fields with the uniqueness property, then every ultraproduct $\prod_{i \in I} F_i/D$ has the uniqueness property.*

Proof. Let F be a valued field. It follows from Krull [10, Th. 19], that there is always some extension of ord to a finite normal extension W of F and that all such extensions of ord are conjugate. To prove that F has the uniqueness property is therefore equivalent to showing that if ord is extended to a finite normal extension W of F and if α, $\alpha' \in W$ are F-conjugate, then $\operatorname{ord}(\alpha) = \operatorname{ord}(\alpha')$.

We may now reformulate the uniqueness property as an intrinsic property of F. First suppose $\operatorname{ord}(F) = Z$. In this case the theory of Newton Polygons as formulated in Dwork [5] shows that the ord values of the roots of a polynomial $J(X) \in F[X]$ are determined by the ord values of the coefficients of $J(X)$. In particular, it follows from the Corollary to Lemma 1.6 of [5] that the roots of $J(X)$ all have the same ord value if and only if the Newton Polygon of $J(X)$ has at most one non-vertical side. Letting $J(X) = \sum_{j=0}^{n} a_j X^j$ and assuming $a_n \neq 0$, $a_0 = 1$, we may formulate the last condition as

$$(*) \qquad j \operatorname{ord}(a_n) \leqq n \operatorname{ord}(a_j), \qquad 0 \leqq j \leqq n.$$

We thereby see that if $\operatorname{ord}(F) = Z$, then the condition that F has the uniqueness property is equivalent to the condition that (*) holds for every irreducible polynomial of degree n with constant term 1. Now the theory of Newton Polygons can be extended to fields with arbitrary value group G and this equivalence remains valid. The proof in Dwork [5] may be adapted in a straightforward manner to a proof of the above equivalence in the general situation.

Now let $\sum_{j=0}^{n} a_j^* X^j$ be a polynomial over $\prod_{i \in I} F_i/D$, irreducible with $a_n^* \neq 0$ and $a_0^* = 1$. Then $\sum_{j=0}^{n} a_j(i) X^j$ is a polynomial over F_i irreducible, $a_n(i) \neq 0$, and $a_0(i) = 1$ for almost all $i \in I$. Thus,

$$j \operatorname{ord}(a_n) \leqq n \operatorname{ord}(a_j) \text{ [a. e.]}, \qquad 0 \leqq j \leqq n,$$

i.e. $j \operatorname{ord}(a_n^*) \leqq n \operatorname{ord}(a_j^*)$, $0 \leqq j \leqq n$, so that $\prod_{i \in I} F_i/D$ has the uniqueness property.

We shall employ the notion of pseudo-convergence introduced by Ostrowski [14]. A detailed investigation of this notion may be found in Kaplansky [8]. Let F be a valued field, and λ a limit ordinal. A sequence $\{a_\rho\}_{\rho<\lambda}$ of elements of F is called *λ-pseudo-Cauchy* if

$$\operatorname{ord}(a_\sigma - a_\rho) < \operatorname{ord}(a_\tau - a_\sigma),$$

for all $\rho < \sigma < \tau < \lambda$.

If $\{a_\rho\}$ is a λ-pseudo-Cauchy sequence, then

$$\operatorname{ord}(a_\sigma - a_\rho) = \operatorname{ord}(a_{\rho+1} - a_\rho),$$

for all $\rho < \sigma < \lambda$.

A sequence $\{a\}_\rho$ is called *pseudo-Cauchy* if it is λ-pseudo-Cauchy for some limit ordinal λ. An element a of F is said to be a *pseudo-limit* of the pseudo-Cauchy sequence $\{a_\rho\}_{\rho<\lambda}$ if $\operatorname{ord}(a - a_\rho) = \operatorname{ord}(a_{\rho+1} - a_\rho)$, for all $\rho < \lambda$. A field F is *λ-pseudo-complete* if it contains a pseudo-limit of each λ-pseudo-Cauchy sequence in F.

In the lemma which follows we give an analytic property of non-principal ultraproducts of valued fields without assuming this property to hold on the components.

LEMMA 9. *Let $\{F_i \mid i \in I\}$ be a denumerably infinite family of valued fields with Z-groups as value groups. Then, every non-principal ultraproduct $\prod_{i \in I} F_i/D$ is ω-pseudo-complete.*

Proof. Let $\{q_k^*\}$ be an ω-pseudo-Cauchy sequence in $\prod_{i\in I} F_i/D$. We may assume that $\{q_k(i)\}$ is an ω-pseudo-Cauchy sequence in F_i, for every $i \in I$. To see this set

$$\gamma_k(i) = \operatorname{ord}(q_{k+1}(i) - q_k(i)), \quad i \in I,\ k < \omega.$$

Then

$$(*) \qquad \gamma_k^* < \gamma_{k+1}^*, \qquad k < \omega.$$

Now define, for all $i \in I$,

$$\delta_0(i) = \gamma_0(i)$$
$$\delta_{k+1}(i) = \max(\delta_k(i) + 1, \gamma_{k+1}(i)).$$

We see inductively from $(*)$ that $\delta_k = \gamma_k$ [a. e.], i. e. $\delta_k^* = \gamma_k^*$, and that $\delta_k(i) < \delta_{k+1}(i)$, for all $i \in I$, $k < \omega$. Next we define, for all $i \in I$,

$$t_0(i) = q_0(i)$$
$$t_{k+1}(i) = \begin{cases} q_{k+1}(i) & \text{if } \operatorname{ord}(q_{k+1}(i) - t_k(i)) = \delta_k(i) \\ t_k(i) + w_k(i) & \text{otherwise.} \end{cases}$$

Here $w_k(i)$ is any element of F_i such that $\operatorname{ord}(w_k(i)) = \delta_k(i)$. Since $\delta_k = \gamma_k$ [a. e.] we see inductively that $t_k = q_k$ [a. e.], i. e. $t_k^* = q_k^*$, and that $\operatorname{ord}(t_{k+1}(i) - t_k(i)) = \delta_k(i)$ for all $i \in \mathrm{I}$, $k < \omega$. We have thus shown that we may assume $\{q_k(i)\}$ to be an ω-pseudo-Cauchy sequence for all $i \in I$.

We may assume that I is the set of non-negative integers. We define the function f by setting $f(i) = q_i(i)$ for all $i < \omega$. We claim that f^* is a pseudo-limit of the sequence $\{q_k^*\}$. Now

$$\operatorname{ord}(q_i(i) - q_k(i)) = \operatorname{ord}(q_{k+1}(i) - q_k(i)),$$

for all $i > k$ since the sequence $\{q_k(i)\}_{k<\omega}$ is pseudo-Cauchy. Hence, for all $k < \omega$

$$\operatorname{ord}(f - q_k) = \operatorname{ord}(q_{k+1} - q_k) \quad \text{[a. e.]}$$

and our claim is established.

Note that the only property of the value groups which is used in the proof is that there is a smallest positive element.

2. *L*-fields. In this section we define L-fields and derive some of their properties. Our interest in L-fields lies in the fact (proved in Section 3) that for any non-principal ultrafilter D on the set P of primes, $\mathfrak{Q} = \prod_{p\in P} Q_p/D$ and $\mathcal{S} = \prod_{p\in P} S_p/D$ are L-fields. The results of this section are used in

Section 3 to prove that $\mathfrak{Q} \approx \mathfrak{S}$, assuming the continuum hypothesis. As we shall see, this isomorphism is equivalent to the principle stated in the introduction.

Definition. An *L-field* is a 4-tuple (V, C, G, ρ) where

(i) V is a valued field of characteristic 0 with valuation group G

(ii) G is a Z-group

(iii) ρ is a prime element of V $(\operatorname{ord}(\rho) = 1)$

(iv) C is a subfield of V such that ord is trivial on C and $\bar{C} = \bar{V}$.

(v) there is an algebraic closure $\tilde{V}$ of V such that

$$\tilde{V} = V(\beta, \rho^{1/k} \colon \beta \in \tilde{C}; k = 1, 2, \cdots).$$

We shall often speak of the L-field V where there is no danger of confusion. By condition (iv) we may identify C with $\bar{V}$. Thus we shall assume $\bar{V} = \bar{C} = C$. If (V, C, G, ρ) and (V', C', G', ρ') are L-fields, an *analytic isomorphism* $\psi \colon V \to V'$ is an algebraic isomorphism of V onto V' such that $\psi(C) = C'$, $\psi(\rho) = \rho'$ and $\psi(\operatorname{ord}(v)) = \operatorname{ord}(\psi(v))$ for all $v \in V$. In case $C = C'$, it is assumed that $\psi \mid C$ is the identity map.

Throughout this section, (V, C, G, ρ) is an L-field. C is a maximal subfield of V on which ord is trivial. We assume that ord has been extended to a valuation of an algebraic closure $\tilde{V}$ of V. If F is a subfield of V, $\tilde{F}$ is the algebraic closure of F, taken in $\tilde{V}$.

LEMMA 10. *Let K be a relatively algebraically closed subfield of a field W of characteristic 0. Let $L \subseteq \tilde{W}$ be such that $[L \colon K] < \infty$. Then*

(i) $[WL \colon W] = [L \colon K]$;

(ii) *L is algebraically closed in WL.*

Proof. Let $\tau \in L$ be such that $L = K(\tau)$. Let A be the monic irreducible polynomial for τ over K. If

$$[W(\tau) : W] < [K(\tau) : K],$$

then A factors over W. However, the coefficients of an irreducible factor of A over W are rational integral polynomials in the conjugates of τ over K. Since the conjugates of τ are all algebraic over K, these coefficients are also algebraic over K and thus are already in K. This proves (i).

Now let $\sigma \in W(\tau)$ be algebraic over $L = K(\tau)$. Using (i), we have

$$\begin{aligned} 1 &= [W(\tau, \sigma) : W(\tau)] = [W(\tau, \sigma) : W]/[W(\tau) : W] \\ &= [K(\tau, \sigma) : K]/[K(\tau) : K], \end{aligned}$$

which establishes (ii).

LEMMA 11. *V possesses the uniqueness property.*

Proof. As in Lemma 8 it suffices to show that if ord is extended to a finite normal extension W of V and if α, $\alpha' \in W$ are V-conjugate, then $\operatorname{ord}(\alpha) = \operatorname{ord}(\alpha')$. By condition (v) for the L-field V, we may suppose (by taking an extension of W, if necessary) that $W = V(\rho^{1/e}, \alpha)$ where $\rho^{1/e}$ is some e-th root of ρ, and that $C(\alpha)$ is a finite normal extension of C containing all e-th roots of unity. Now C is algebraically closed in V; for otherwise there exists $\beta \in V - C$ which is algebraic over C. But then ord is trivial on $C(\beta) \supsetneq C$, a contradiction.

By Lemma 10, $[V(\alpha) : V] = [C(\alpha) : C]$ ($= f$, say). Thus $1, \bar{\alpha}, \cdots, \bar{\alpha}^{f-1}$ form a basis for $\bar{W}/\bar{V}$. Also $1, \rho^{1/e}, \cdots, (\rho^{1/e})^{e-1}$ represent distinct cosets for $\operatorname{ord}(W)$ modulo G. Let $\eta_{ij} = (\rho^{1/e})^i\alpha^j$. A standard argument shows that for all $v_{ij} \in V$,

$$(*) \qquad \operatorname{ord}\Big(\sum_{i=0}^{e-1}\sum_{j=0}^{f-1} v_{ij}\eta_{ij}\Big) = \min_{i,j}[\operatorname{ord}(v_{ij}\eta_{ij})].$$

Since the η_{ij} span W as a V-linear space, the η_{ij} form an integral basis for W/V.

Now let σ be a V-automorphism of W. Thus $\sigma(\sum_{i,j} v_{ij}\eta_{ij}) = \sum_{i,j} v_{ij}\sigma(\eta_{ij})$. Now $\sigma(\eta_{ij}) = \sigma(\rho^{1/e})^i\sigma(\alpha)^j$. But $\sigma(\rho^{1/e})$ is an e-th root of ρ and $C(\sigma(\alpha)) = C(\alpha)$ is a finite normal extension of C containing all e-th roots of unity. Thus again

$$(**) \qquad \operatorname{ord}(\sigma(\sum_{i,j} v_{ij}\eta_{ij})) = \operatorname{ord}(\sum_{i,j} v_{ij}\sigma(\eta_{ij})) = \min_{i,j}[\operatorname{ord}(v_{ij}\sigma(\eta_{ij})].$$

We have $\operatorname{ord}(\sigma(\eta_{ij})) = \mathrm{i}(1/e) = \operatorname{ord}(\eta_{ij})$ where $1/e$ is the unique element γ of $\operatorname{ord}(W)$ such that $e\gamma = 1$. Using (*) and (**) this completes the proof of Lemma 11.

LEMMA 12. *Let F be a valued field of characteristic 0. If F satisfies the uniqueness property then F satisfies the Hensel-Rychlik property.*

Proof. Let $J \in \mathfrak{O}_F[X]$ be monic, with no multiple roots. Assume $f \in \mathfrak{O}_F$ is such that $\operatorname{ord}(J(f)) > \operatorname{ord}(D(J))$.

Let $J(X) = \prod_{i=1}^{n}(X - \alpha_i)$ be the splitting of J over $\bar{F}$. Thus

$$(*) \qquad \sum_i \operatorname{ord}(f - \alpha_i) > \sum_{i \neq j} \operatorname{ord}(\alpha_i - \alpha_j).$$

This implies that there exists a root α_k of J such that

$$(**) \qquad \operatorname{ord}(f - \alpha_k) > \operatorname{ord}(\alpha_i - \alpha_k),\ i \neq k.$$

Indeed, otherwise for each k there would be an $m(k) \neq k$ such that

$$\operatorname{ord}(f-\alpha_k) \leqq \operatorname{ord}(\alpha_{m(k)}-\alpha_k).$$

This contradiction to (*) establishes (**). We may assume $k=1$.

We now show that $\alpha_1 \in \mathfrak{O}_F$; this will prove the Hensel-Rychlik property for F. Suppose to the contrary that $F(\alpha_1) \supsetneq F$. Let σ be a non-trivial F-monomorphism of $F(\alpha_1)$ into $\tilde{F}$. There exists an $i>1$ such that $\sigma(\alpha_1)=\alpha_i$. This shows

$$\sigma(f-\alpha_1)=f-\alpha_i.$$

By hypothesis, $\operatorname{ord}(\sigma(f-\alpha_1))=\operatorname{ord}(f-\alpha_1)$. Thus $\operatorname{ord}(f-\alpha_1)=\operatorname{ord}(f-\alpha_i)$. This contradicts (**) and completes the proof of the lemma.

Combining Lemma 11 and Lemma 12 we obtain the following frequently used fact.

LEMMA 13. *V satisfies the Hensel-Rychlik property.*

LEMMA 14. *If V_0 is a relatively algebraically closed subfield of V such that $V_0 \supseteq C(\rho)$, then V_0 (or more precisely $(V_0, C, \operatorname{ord}(V_0), \rho)$) is an L-field.*

Proof. We first show $G_0=\operatorname{ord}(V_0)$ is a pure subgroup of G. Suppose $\gamma \in G$, $0<n \in Z$, and $n\gamma \in G_0$. Let $a \in V$ be such that $\operatorname{ord}(a)=\gamma$, and let $b \in V_0$ be such that $\operatorname{ord}(b)=n\gamma$. Thus there exists a $u \in V$ such that $a^n=bu$ and $\operatorname{ord}(u)=0$. Let $c \in C$ be such that $\operatorname{ord}(cu-1)>0$. By replacing b by bc^{-1}, we may assume $\operatorname{ord}(u-1)>0$. By Lemma 13, V has the Hensel-Rychlik property. Thus there exists $v \in V$ such that $u=v^n$. By replacing a by av^{-1}, we may assume $a^n=b$. Since V_0 is algebraically closed in V, $a \in V_0$. Thus $\gamma=\operatorname{ord}(a) \in G_0$. This shows that G_0 is a pure subgroup of G containing 1 as a member. It follows that G_0 is a Z-group, which proves V_0 satisfies condition (ii) for an L-field.

We now demonstrate that V_0 satisfies condition (v) for an L-field. Let $\alpha \in \tilde{V}_0 \subseteq \tilde{V}$. Then there exists a positive integer k and $\beta \in \tilde{C}$ such that $\alpha \in V(\beta, \rho^{1/k})$. By Lemma 10, $\alpha \in V_0(\beta, \rho^{1/k})$. This shows that V_0 satisfies condition (v). V_0 also satisfies conditions (i), (iii), and (iv). This proves the lemma.

PROPOSITION 1. *Let F be a relatively algebraically closed subfield of V. If $\alpha \in \tilde{F}$ and $y \in V$, then there exists $\phi \in F$ such that* $\operatorname{ord}(y-\phi) \geqq \operatorname{ord}(y-\alpha)$.

Proof. By condition (v) for the L-field V, there is a positive integer k and a $\beta \in \tilde{C} \subset \tilde{V}$ such that $\alpha \in L=V(\rho^{1/k}, \beta)$, where $\rho^{1/k}$ is some k-th root of ρ.

Let $m=[V(\beta):V]$, and set $\eta_{ij}=(\rho^{1/k})^i\beta^j$, $(i=0,\cdots,k-1,$

$j=0,\cdots,m-1$). As in the proof of Lemma 11, the η_{ij} form an integral basis for L/V.

In particular, we have

$$\operatorname{ord}\Big(\sum_{i=0}^{k-1}\sum_{j=0}^{m-1} v_{ij}\eta_{ij}\Big)=\min_{i,j}(\operatorname{ord}(v_{ij}\eta_{ij})).$$

We know $\alpha\in L=VF(\rho^{1/k},\beta)$. By Lemma 10, $\alpha\in F(\rho^{1/k},\beta)$. Since the η_{ij} form a basis for $F(\rho^{1/k},\beta)/F$, there exist $f_{ij}\in F$ such that $\alpha=\sum_{i,j} f_{ij}\eta_{ij}$. Defining

$$d_{ij}=\begin{cases}1 \text{ if } i=j=0\\ 0 \text{ otherwise}\end{cases}$$

we have

$$\alpha-y=\sum_{i,j}(f_{ij}-d_{ij}y)\eta_{ij}$$

Now

$$\begin{aligned}\operatorname{ord}(\alpha-y)&=\min_{i,j}(\operatorname{ord}((f_{ij}-d_{ij}y)\eta_{ij})\\ &\leqq \operatorname{ord}(f_{00}-d_{00}y)=\operatorname{ord}(f_{00}-y).\end{aligned}$$

Thus $\phi=f_{00}$ is an element of F with the required property.

We recall some definitions and properties of general valued fields E, F. All these facts, including Lemma 15 and Lemma 16, may be found in Kaplansky [8]. If $F\supset E$, $\operatorname{ord}(F)=\operatorname{ord}(E)$, and $\bar F=\bar E$, then F is an *immediate* extension of E. E is *maximally complete* if it has no immediate extensions.

LEMMA 15 [8, Th. 4]. *E is maximally complete if and only if it contains a pseudo-limit of each pseudo-Cauchy sequence in E.*

LEMMA 16 [8, Th. 5]. *If the characteristic of $\bar E$ is 0, then there exists a unique (up to E-analytic isomorphism) immediate, maximally complete extension of E.*

Definition. A field F is *V-maximally complete* if $C(\rho)\subseteq F\subseteq V$, and there is no immediate extension of F in V.

Remark. If F is V-maximally complete and $\operatorname{ord}(F)$ is pure in V, then F is relatively algebraically closed in V.

PROPOSITION 2. *Assume V is ω-pseudo complete. If F is V-maximally complete and* $\operatorname{ord}(F)=H$ *is a pure countable subgroup of G, then F is maximally complete.*

Proof. We need only show that every pseudo-Cauchy sequence in F has a pseudo-limit in F. Since $\operatorname{ord}(F)$ is countable, every pseudo-Cauchy sequence is countable, and hence has a cofinal ω-pseudo-Cauchy subsequence.

It thus suffices to prove F is ω-pseudo-complete. Let the sequence $\{f_i\}_{i<\omega}$ in F be pseudo-Cauchy. Since V is ω-pseudo-complete, $\{f_i\}$ has a pseudo-limit $q \in V$. First suppose that there exists $f \in F$ such that

$$(*) \qquad \operatorname{ord}(q-f) > \operatorname{ord}(q-f_i), \; i < \omega.$$

Then $\operatorname{ord}(f-f_i) = \operatorname{ord}((f-q)+(q-f_i)) = \operatorname{ord}(q-f_i)$, $i < \omega$, which implies that f is a pseudo-limit of $\{f_i\}$. Suppose now that there does not exist $f \in F$ such that $(*)$ holds. We then show that q is already in F, by showing $F(q)/F$ is immediate. It suffices to show that for all $A \in F[X]$ we have $\operatorname{ord}(A(q)) =$ eventual value of $\operatorname{ord}(A(f_i))$. By factoring A over $\tilde{F}$, it suffices to show that for all $\alpha \in \tilde{F}$, $\operatorname{ord}(q-\alpha) =$ eventual value of $\operatorname{ord}(f_i-\alpha)$. Our assumption that H is pure in G guarantees by the preceding Remark that F is relatively algebraically closed in V. Hence Proposition 1 applies and shows that there exists $f \in F$ such that $\operatorname{ord}(q-f) \geqq \operatorname{ord}(q-\alpha)$. Now there exists $i_0 < \omega$ such that if $i > i_0$ then

$$\operatorname{ord}(q-f_i) > \operatorname{ord}(q-f_{i-1}) \geqq \operatorname{ord}(q-f) \geqq \operatorname{ord}(q-\alpha).$$

Thus if $i > i_0$,

$$\operatorname{ord}(f_i-\alpha) = \min[\operatorname{ord}(q-f_i), \operatorname{ord}(q-\alpha)] = \operatorname{ord}(q-\alpha).$$

This completes the proof of Proposition 2.

Definition. If H is a subgroup of an abelian group K, then $(H \mid K)$ denotes the pure subgroup of K consisting of all $\alpha \in K$ such that there exists a positive integer m for which $m\alpha \in H$.

Definition. If $E \supseteq C(\rho)$ is a subfield of V, we denote by $[E \mid V]$ some fixed subfield L of V, maximal with respect to the property:

$$\operatorname{ord}(L) \subseteq (\operatorname{ord}\ (E) \mid G).$$

Remark. The existence of an $[E \mid V]$ is guaranteed by Zorn's Lemma. $[E \mid V]$ is V-maximally complete. By the Remark preceding Lemma 3, $[E \mid V]$ is relatively algebraically closed in V.

PROPOSITION 3. *Let (V, C, G, ρ) and (V', C, G, ρ') be ω-pseudo-complete L-fields with common sub-field C and group G. Assume that there exists a cross-section $\pi\colon G \to V$ (resp.: $\pi'\colon G \to V'$) such that $\pi(1) = \rho$ (resp.: $\pi'(1) = \rho'$). Let $F \supseteq C(\rho)$ (resp.: $F' \supseteq C(\rho')$) be V-maximally complete (resp.: V'-maximally complete), and suppose $\operatorname{ord}(F) = \operatorname{ord}(F')$ is a countable, pure subgroup of G. Then $(F, C, \operatorname{ord}(F), \rho)$ and $(F', C, \operatorname{ord}(F'), \rho')$ are L-fields. Moreover, if $\psi\colon F \to F'$ is an analytic isomorphism of L-fields, and $x \in V - F$, then there exists an extension ψ_0 of ψ to an analytic isomorphism of the L-field $[F(x) \mid V]$ onto some sub-L-field F_1' of V'. In addition, $\operatorname{ord}([F(x) \mid V]) = (\operatorname{ord}(F(x)) \mid G) = \operatorname{ord}(F_1')$.*

Proof. The Remark preceding Proposition 2 together with Lemma 14 show that F (resp.: F') is an L-field. The V-maximal completeness of F implies that $\operatorname{ord}(F(x)) \supsetneqq \operatorname{ord}(F)$. Thus there exists $A \in F[X]$ such that $\operatorname{ord}(A(x)) \notin \operatorname{ord}(F)$. As $\operatorname{ord}(F)$ is a pure subgroup of G, this relation is equivalent to $\operatorname{ord}(A(x)) \in G-(\operatorname{ord}(F) \mid G)$, which in turn implies $\operatorname{ord}(A(x)) \notin \operatorname{ord}(\tilde{F})$. By factoring A over $\tilde{F}$, we see there exists an $\alpha \in \tilde{F}$ such that $\operatorname{ord}(x-\alpha) \notin \operatorname{ord}(\tilde{F})$. By Proposition 1, there is an $f \in F$ such that $\operatorname{ord}(x-f) \geqq \operatorname{ord}(x-\alpha)$. If $\operatorname{ord}(x-f) > \operatorname{ord}(x-\alpha)$, then $\operatorname{ord}(f-\alpha) = \operatorname{ord}(x-\alpha) \notin \operatorname{ord}(\tilde{F})$, a contradiction. Thus

$$\operatorname{ord}(x-f) = \operatorname{ord}(x-\alpha) \in G-(\operatorname{ord}(F) \mid G).$$

Replacing x by $x-f$, we may assume

$$\gamma = \operatorname{ord}(x) \in G-(\operatorname{ord}(F) \mid G) = G-\operatorname{ord}(F).$$

We regard C as a common subfield of V and V'.

Since $\operatorname{ord}(x\pi(-\gamma)) = 0$, there exists $c \in C \subseteq F$ such that

$$\operatorname{ord}(cx\pi(-\gamma)-1) > 0.$$

Replacing x by cx, we may assume $\operatorname{ord}(x\pi(-\gamma)-1) > 0$. We set $x' = \pi'(-\gamma)$. Then there exists a unique analytic isomorphism $\psi_1 \colon F(x) \to F'(x')$ which extends ψ and such that $\psi_1(x) = x'$. Indeed, since $\gamma \notin (\operatorname{ord}(F) \mid G)$, x (resp.: x') is transcendental over F (resp.: F'). Thus ψ_1 is an algebraic isomorphism. If $A(X) = \sum a_i X^i \in F[X]$, we have

$$\begin{aligned} \operatorname{ord}(A(x)) &= \min_i (\operatorname{ord}(a_i) + i\gamma) \\ &= \min_i (\operatorname{ord}(\psi(a_i)) + i\gamma) = \operatorname{ord}(\textstyle\sum \psi(a_i) x'^i) \\ &= \operatorname{ord} \psi_1(A(x)), \end{aligned}$$

so ψ_1 is an analytic isomorphism.

Since G is a Z-group, for all positive integers n, there exist unique $\gamma(n) \in G$ and $r(n) \in Z$ such that

$$\gamma = n\gamma(n) + r(n), 0 \leqq r(n) < n.$$

We define $y_n = x\pi(-r(n)) = x\rho^{-r(n)} \in F(x)$

$$(\text{resp.: } y_n' = \pi'(n\gamma(\mathrm{n})) = x'(\rho')^{-r(n)} = \psi_1(y_n) \in F'(x')).$$

Thus $\operatorname{ord}(y_n\pi(-\gamma(n))^n - 1) = \operatorname{ord}(x\pi(-\gamma)-1) > 0$. By the Hensel-Rychlik property, $y_n\pi(-\gamma(n))^n$ has an n-th root in V. Thus y_n has an n-th root z_n in V and hence in $[F(x) \mid V]$. The field $F(z_n) \supseteq F(x)$ is independent of which n-th root of y_n is chosen since all roots of unity in V are

already in $C \subseteq F$. We let $z_n' = \pi'(\gamma(n))$. Thus z_n' is an n-th root of y_n'. Hence, $z_n' \in [F'(x') \mid V']$.

We assert that, for each positive integer n, there exists an extension of ψ_1 to an analytic isomorphism $\sigma\colon F(z_n) \to F'(z_n')$ defined by requiring $\sigma(z_n) = z_n'$. σ is an algebraic isomorphism because $z_n{}^n = y_n$, $z_n'{}^n = y_n' = \psi_1(y_n)$ and, as we will now see, $X^n - y_n$ is irreducible over $F(x)$. To see that $X^n - y_n$ is irreducible over $F(x)$, it suffices to show that

$$(\operatorname{ord}(F(z_n)) : \operatorname{ord}(F(x))) \geqq n, \text{ i.e. } 0 < j < n$$

implies $j\gamma(n) \notin \operatorname{ord}(F(x))$. This will also show that σ is analytic. Assume $m \in Z$ and $\delta \in \operatorname{ord}(F)$ are such that

$$j\gamma(n) = \delta + m\gamma \in \operatorname{ord}(F(x)).$$

Then $j\gamma(n) = \delta + m(n\gamma(n) + r(n)) = mn\gamma(n) + \delta + mr(n)$. Thus

$$(j - mn)\gamma(n) = \delta + mr(n) \in \operatorname{ord}(F).$$

It follows that

$$\begin{aligned}(j - mn)\gamma &= (j - mn)(n\gamma(n) + r(n)) \\ &= n(j - mn)\gamma(n) + (j - mn)r(n) \in \operatorname{ord}(F).\end{aligned}$$

Since $0 < j < n$, we have $\gamma \in (\operatorname{ord}(F) \mid G)$, a contradiction. This proves our assertion that the set Σ_n of analytic isomorphisms $\sigma\colon F(z_n) \to F'(z_n')$ extending ψ_1 is non-empty.

Thus $1 \leqq |\Sigma_n| \leqq n$. We now assert there exists an extension of ψ_1 to an analytic isomorphism

$$\psi_2\colon F(z_n\colon n = 1, 2, \cdots) \to F'(z_n'\colon n = 1, 2, \cdots).$$

We give each Σ_n the discrete topology and $\Sigma = \prod_{n=1}^{\infty} \Sigma_n$, the product topology. By Tychonoff's Theorem, Σ is compact. For each natural number n, let $T(n)$ be the (closed) set of $t \in \Sigma$ such that $t(m) = t(n) \mid F(z_m)$ if $F(z_m) \subseteq F(z_n)$. Suppose $d \mid n$, say $n = de$. Then $(z_n{}^e)^d = y_n$. Let a, $r \in Z$ be such that

$$r(n) = ad + r,\ 0 \leqq r < d.$$

Thus $\gamma = n\gamma(n) + r(n) = n\gamma(n) + ad + r = d(e\gamma(n) + a) + r$. This implies $r = r(d) = r(n) - ad$. It follows that

$$(z_n{}^e)^d = x\pi(-r(n)) = x\pi(-r(d) - ad).$$

Hence $(z_n{}^e\pi(a))^d = x\pi(-r(d)) = y_d$. Thus $F(z_n) \supseteq F(z_d)$. This implies that $T(n) \cap T(n') \supseteq T(nn')$ for all positive integers n and n', so that

the $T(n)$ have the finite intersection property. Thus $\bigcap_{n=1}^{\infty} T(n) \neq \phi$. Any $t \in \bigcap_{n=1}^{\infty} T(n)$ defines an extension of ψ_1 to an analytic isomorphism

$$\psi_2 \colon F(z_n \colon 1, 2, \cdots) \to F'(z_n' \colon n = 1, 2, \cdots)$$

by requiring that

$$\psi_2 \mid F(z_n) = t(n).$$

We next claim that $\operatorname{ord}(F(z_n \colon 1, 2, \cdots)) = (\operatorname{ord}(F(x)) \mid G)$. Let $\delta \in (\operatorname{ord}(F(x)) \mid G)$ and let m be a positive integer such that $m\delta \in \operatorname{ord}(F(x))$, say

$$m\delta = h + n\gamma \text{ with } h \in \operatorname{ord}(F), \text{ and } n \in Z.$$

Let $h_1 \in G$ and $m_1 \in Z$ be such that

$$h = mh_1 + m_1,\ 0 \leqq m_1 < m.$$

Since $\operatorname{ord}(F)$ is a pure subgroup of G and $h - m_1 \in \operatorname{ord}(F)$, we see $h_1 \in \operatorname{ord}(F)$. We have

$$\begin{aligned} m\delta &= mh_1 + m_1 + n\delta = mh_1 + m_1 + n(m\delta(m) + r(m)) \\ &= m(h_1 + n\gamma(m)) + m_1 + nr(m). \end{aligned}$$

It follows that $m \mid (m_1 + nr(m))$ and

$$\delta = h_1 + n\gamma(m) + (m_1 + nr(m))/m.$$

Thus $\delta \in \operatorname{ord}(F(z_n \colon n = 1, 2, \cdots))$. This shows $\operatorname{ord}(F(z_n \colon n = 1, 2, \cdots)) \supseteq (\operatorname{ord}(F(x)) \mid G)$. The reverse inclusion is a direct consequence of the definitions. This establishes our claim.

We may now conclude that $\operatorname{ord}([F(x) \mid V]) = (\operatorname{ord}(F(x) \mid G)$ and that $[F(x) \mid V]$ (resp. $[F'(x') \mid V']$) is an immediate extension of $F(z_n \colon n = 1, 2, \cdots)$ (resp.: $F'(z_n' \colon n = 1, 2, \cdots)$). As $\operatorname{ord}(F)$ is countable, so is $(\operatorname{ord}(F(x)) \mid G)$. It follows from Proposition 2 that $[F(x) \mid V]$ (resp.: $[F'(x') \mid V]$) is maximally complete. By Lemma 16, there exists an extension ψ_0 of ψ_2,

$$\psi_0 \colon [F(x) \mid V] \to [F'(x') \mid V'] = F_1'$$

which is an analytic isomorphism.

This completes the proof of Proposition 3.

THEOREM 2. *Let V and V' be two ω-pseudo-complete fields of cardinality $\aleph_1$ with the same value group G of cardinality $\aleph_1$ and isomorphic residue class fields.*

Assume that there exists a cross-section $\pi\colon G \to V$ (*resp.*: $\pi'\colon G \to V'$) *such that* $\pi(1) = \rho$ (*resp.*: $\pi'(1) = \rho'$). *Then* V *and* V' *are analytically isomorphic.*

Remark. We may regard the residue class field C as a common subfield of V and V'.

Proof. Choose a transcendence base B (resp.: B') of V (resp.: V') over $C(\rho)$ (resp.: $C(\rho')$). Since $|G| = \aleph_1$, it immediately follows that $|B| = |B'| = \aleph_1$. We may assume that the prime elements ρ (resp.: ρ') of V (resp.: V') lie in B (resp.: B'). Well-order B and B' to form ω_1-sequences (ω_1 being the first uncountable ordinal):

$$B = \{x_1, \cdots, x_\eta, \cdots\}\eta < \omega_1$$
$$B' = \{x_1', \cdots, x_\eta', \cdots\}\eta < \omega_1.$$

We are now going to define, for $0 < \lambda < \omega_1$, a V-maximally complete field F_λ such that $\operatorname{ord}(F_\lambda)$ is a pure, countable subgroup of G, a V'-maximally complete field F_λ', and an analytic isomorphism $\psi_\lambda\colon F_\lambda \to F_\lambda'$.

For $\lambda = 1$, we set $F_1 = [C(\rho) \mid V]$, $F_1' = [C(\rho') \mid V']$. We have an analytic isomorphism $\phi\colon C(\rho) \to C(\rho')$. Since $\operatorname{ord}(C(\rho)) = Z$ is a pure, countable subgroup of G we see that $[C(\rho) \mid V]$ (resp.: $[C(\rho') \mid V']$) is a maximally complete immediate extension of $C(\rho)$ (resp.: $C(\rho')$) and hence ϕ extends to an analytic isomorphism $\psi_1\colon F_1 \to F_1'$, by Lemma 16.

Suppose $\lambda < \omega_1$ is an odd ordinal (i.e., λ is the sum of a limit ordinal and an odd integer), and we have defined $\psi_\lambda\colon F_\lambda \to F_\lambda'$. Since $\operatorname{ord}(F_\lambda)$ is countable, it follows easily that there exists a $\mu < \omega_1$ such that $x_\mu \notin F_\lambda$. Let $\mu(\lambda)$ be the smallest such μ. Set $F_{\lambda+1} = [F(x_{\mu(\lambda)}) \mid V]$. By Proposition 3, there exist a subfield $F'_{\lambda+1} \subset V'$ and an analytic isomorphism $\psi_{\lambda+1}\colon F_{\lambda+1} \to F'_{\lambda+1}$ extending ψ_λ. It follows from Proposition 3, that $\psi_{\lambda+1}$, $F_{\lambda+1}$, and $F'_{\lambda+1}$ satisfy our inductive hypothesis. If λ is an even ordinal, then we proceed in a similar fashion, reversing the roles of F_λ with F'_λ and replacing ψ_λ by ψ_λ^{-1}.

If σ is a limit ordinal $< \omega_1$, we form the fields $D_\sigma = \bigcup_{\lambda<\sigma} F_\lambda$, $D'_\sigma = \bigcup_{\lambda<\sigma} F'_\lambda$. Then clearly there is an analytic isomorphism $\phi\colon D_\sigma \to D'_\sigma$ extending all the ψ_λ, $\lambda < \sigma$. Letting $F_\sigma = [D_\sigma \mid V]$ and $F'_\sigma = [D'_\sigma \mid V']$, we may extend ϕ to an analytic isomorphism $\psi_\sigma\colon F_\sigma \to F'_\sigma$, since F_σ (resp.: F'_σ) is an immediate maximally complete extension of D_σ (resp.: D'_σ).

We have thus inductively proved that the fields $W = \bigcup_{\lambda<\omega_1} F_\lambda$ and $W' = \bigcup_{\lambda<\omega_1} F'_\lambda$ are analytically isomorphic. By alternating the isomorphisms

between F_λ and F'_λ for λ even and odd, we have guaranteed that both B and B' are exhausted. Hence V (resp.: V') is an algebraic extension of the field W (resp.: W'). On the other hand, since each F_λ (resp.: F'_λ) is relatively algebraically closed in V (resp.: V'), it follows that W (resp.: W') is also relatively algebraically closed in V (resp.: V'). Hence $V = W$ and $V' = W'$, and the theorem is proved.

Theorem 2 clearly calls for the identification of ω-pseudo-complete L-fields of cardinality $\aleph_1$ with groups of cardinality $\aleph_1$.

Definition. Let C be a field and G an ordered abelian group. Let α be an infinite cardinal. Then $C(t^G)_\alpha$ denotes the valued field of formal series

$$\sum_{\gamma \in S} c_\gamma t^\gamma,$$

where S is an arbitrary well-ordered subset of G of cardinality at most α, and $c_\gamma \in C$ for all $\gamma \in S$.

The proof that $C(t^G)_\alpha$ forms a valued field is easily generalized from the case $\alpha = |G|$ given in the literature. (See e.g. Schilling [16, p. 23].)

THEOREM 3. *Let V satisfy the hypothesis of Theorem 2. Then V is analytically isomorphic to $\bar{V}(t^G)_{\aleph_0}$, assuming the continuum hypothesis.*

Proof. Since the field $V' = \bar{V}(t^G)_{\aleph_0}$ has cardinality $2^{\aleph_0}$, it suffices by Theorem 2 and the continuum hypothesis to show that V is ω-pseudo-complete. Since every ω-pseudo-complete sequence can be valued in a countable subgroup H of G, this follows from the fact that $\bar{V}(t^H)_{|H|}$ is pseudo-complete [16, p. 23, Corollary to Theorem 8].

3. Applications. For each rational prime p, let x_p be transcendental over R_p and let S_p denote the completion of the field $R_p(x_p)$ with respect to the prime defined by x_p. The valued field S_p may be identified with the field of formal meromorphic expansions about zero over R_p, i.e. series of the form $\sum_{k=-n}^{\infty} \alpha_k x_p^k$ with $\alpha_k \in R_p$. For S_p, as for Q_p, the value group is Z, the group of rational integers with the usual ordering, and $\bar{S}_p = \bar{Q}_p = R_p$. It follows that if D is a non-principal ultafilter on the set P of primes, then $\mathfrak{S} = \prod_{p \in P} S_p/D$ and $\mathfrak{Q} = \prod_{p \in P} Q_p/D$ are valued fields with common valuation group $\mathfrak{Z} = \prod_{p \in P} Z/D$. In order to show that $\mathfrak{Q}$ and $\mathfrak{S}$ are L-fields, we shall need the following fact.

LEMMA 17. *Let V be a field of characteristic 0 with valuation* ord *in a Z-group G. Suppose V satisfies:*

(a) *the solvability property;*

(b) *the uniqueness property.*

Then there exists a subfield C of V and a prime element $\rho \in V$ such that (V, C, G, ρ) is an L-field. Moreover, of C_1 is any subfield of V on which ord *is trivial with $\overline{C_1} = \bar{V}$ and if $\rho_1 \in V$ is such that* $\operatorname{ord}(\rho_1) = 1$, *then (V, C_1, G, ρ_1) is an L-field.*

Proof. By Lemma 12, we have that V satisfies the Hensel-Rychlik property. By Lemma 3, V contains a maximal subfield C on which ord is trivial, with $\bar{C} = \bar{V}$. Now let C_1 be any subfield of V on which ord is trivial with $\overline{C_1} = \bar{V}$ and let $\rho_1 \in V$ be such that $\operatorname{ord}(\rho_1) = 1$. It remains to show $\tilde{V} = V(\beta, \rho_1{}^{1/k}\colon \beta \in \tilde{C}_1;\ k = 1, \cdots)$.

Let $\alpha \in V$. We show that there exist $\rho_1{}^{1/k}$ and $\beta \in \tilde{C}_1$ such that $\alpha \in V(\beta, \rho_1{}^{1/k})$. By condition (a) there exist $a_1, \cdots, a_j \in \tilde{V}$ and positive integers $k_1, \cdots, k_j$ such that $a_1 \in V$, $a_{i+1} \in V(a_1{}^{1/k_1}, \cdots, a_i{}^{1/k_i})$ for $1 \leqq i \leqq j-1$, and $\alpha \in V(a_1{}^{1/k_1}, \cdots, a_j{}^{1/k_j})$ for some choice of the roots. Since conditions (a) and (b) are preserved under finite extensions, we may inductively consider one root adjuction at a time.

Our assertion will therefore hold if for all $a \in V$ and for all positive integers k, $X^k - a$ splits completely in a field $L = V(\rho_1{}^{1/k}, \beta)$. Because $\tilde{C}_1$ contains all roots of unity, it is even sufficient to show that $X^k - a$ has a root in L. Let $\operatorname{ord}(a) = \gamma$. Since G is a Z-group, there exist $\delta \in G$ and $r \in Z$ such that

$$\gamma = k\delta + r, \quad 0 \leqq r < k.$$

Let $b \in V$ be such that $\operatorname{ord}(b) = \delta$. There exists a unique $d \in V$ such that $a = b^k d$; we have $0 \leqq \operatorname{ord}(d) = r < k$. We define u by $d = u\rho_1{}^r$. It follows that $\operatorname{ord}(u) = 0$. Let $c \in C_1$ be such that $\operatorname{ord}(u/c - 1) > 0$. If $v = u/c \in V$, then $\operatorname{ord}(v) = 0$ so that $X^k - \bar{v}$ has no multiple roots in $\bar{V}$. Since V satisfies the Hensel-Rychlik property, there exists $w \in V$ such that $w^k = v$. Then $b\rho_1{}^{r/k} w c^{1/k} \in V((\rho_1{}^r c)^{1/k})$ is a root of $X^k - a$. This completes the proof of the lemma.

LEMMA 18. *The non-principal ultraproducts*

$$\mathfrak{Q} = \prod_{p \in P} Q_p/D \text{ and } \mathfrak{S} = \prod_{p \in P} S_p/D$$

are ω-pseudo-complete L-fields with cross-sections, with common value group $\mathfrak{Z} = \prod_{p \in P} Z/D$ and residue class fields isomorphic to $\mathfrak{R} = \prod_{p \in P} R_p/D$. In addition $|\mathfrak{Z}| = |\mathfrak{Q}| = |\mathfrak{S}| = 2^{\aleph_0}$.

Proof. Lemma 1 shows that $\bar{\mathfrak{Q}} = \bar{\mathfrak{S}} = \mathfrak{R}$. By Lemma 4, the characteristic of the common residue class field is 0. It follows from the Hilbert

Theory [18, Prop. 3-6-6] that Q_p and S_p satisfy the solvability property. Then from Lemma 5, $\mathfrak{Q}$ and $\mathfrak{S}$ satisfy the solvability property. It follows from the completeness of Q_p and S_p that they satisfy the uniqueness property [7, Chap. II, § 12.3]. By Lemma 8, $\mathfrak{Q}$ and $\mathfrak{S}$ possess the uniqueness property. We now see that $\mathfrak{Q}$ and $\mathfrak{S}$ satisfy the hypothesis of Lemma 17. Thus $\mathfrak{Q}$ and $\mathfrak{S}$ are L-fields. The ω-completeness follows from Lemma 9. Lemma 7 proves the existence of the cross-sections. The cardinality assertion is proved in Lemma 6. This completes the proof of the lemma.

We may now apply Theorem 2 to obtain the following isomorphism theorem.

THEOREM 4. *Assuming the continuum hypothesis, $2^{\aleph_0} = \aleph_1$, the ultraproducts $\mathfrak{Q} = \prod_{p\in P} Q_p/D$ and $\mathfrak{S} = \prod_{p\in P} S_p/D$ are analytically isomorphic, for every non-principal ultrafilter D.*

For the purposes of our applications, it is desirable to eliminate the continuum hypothesis. Without the continuum hypothesis the proof of Theorem 4 establishes the following result, which is sufficient for these applications.

THEOREM 4′. *Let D be a non-principal ultrafilter, and let $\mathfrak{Q} = \prod_{p\in P} Q_p/D$ and $\mathfrak{S} = \prod_{p\in P} S_p/D$. Suppose ψ_0 is an analytic isomorphism of subfields $\mathfrak{Q}_0$ and $\mathfrak{S}_0$ of $\mathfrak{Q}$ and $\mathfrak{S}$ respectively. Let $\mathfrak{Q}_1/\mathfrak{Q}_0$ be a sub-extension of $\mathfrak{Q}/\mathfrak{Q}_0$ and suppose that $|\operatorname{ord}(\mathfrak{Q}_1)| = \aleph_0$. Then there exists an extension of ψ_0 to an analytic monomorphism of $\mathfrak{Q}_1$ into $\mathfrak{S}$. Of course, the roles of $\mathfrak{Q}$ and $\mathfrak{S}$ may be reversed.*

THEOREM 5. *Given a set $d_1,\cdots,d_r$ of positive rational integers, there exists a finite set $A = A(d_1,\cdots,d_r)$ of primes such that for all primes $p \notin A$ every system $f_1,\cdots,f_r$ of homogeneous polynomials over Q_p, with degree $f_i = d_i$, $1 \leqq i \leqq r$, and with $n > \sum_{i=1}^{r} d_i^2$ variables has a common non-trivial zero in Q_p.*[3]

Proof. The analogous assertion for the S_p (with A empty) is proved using the theorem of Chevalley mentioned in the proof of Theorem 1.[4] Now assume, contrary to the theorem, that there is an infinite subset S of P such that for all $p \in S$ there exist homogeneous polynomials

$$f_i = \sum a^{(i)}{}_{j_1\cdots j_n}(p) X_1^{j_1}\cdots X_n^{j_n}, \qquad i = 1,\cdots,r$$

[3] In [3] we show that the set $A(d_1,\cdots,d_r)$ is computable (in the sense of recursive function theory) from the degrees $d_1,\cdots,d_r$.

[4] See Lang [7], Corollary to Theorems 6 and 8.

of degree d_i over Q_p with no non-trivial solution over Q_p. Let D be a non-principal ultrafilter on P containing S as a member. For $p \in P - S$, and for $0 < j_1 + \cdots + j_n \leqq d_i$, $i = 1, \cdots, r$, we set $a^{(1)}{}_{j_1 \cdots j_n}(p) = 0$. We now set

$$f_i^* = \sum a^*{}_{j_1 \cdots j_n} X_1{}^{j_1} \cdots X_n{}^{j_n} \qquad j_1 + \cdots + j_n = d_i.$$

Then $f_1^* = \cdots = f_r^* = 0$ is a system of polynomial equations over $\mathfrak{Q}$ in n variables without a non-trivial solution over $\mathfrak{Q}$. We have $d_i =$ degree f_i^*. Now assume the continuum hypothesis. Then $\mathfrak{Q} \approx \mathcal{S}$, so that there exists a similar system of equations over $\mathcal{S}$ without a non-trivial solution over $\mathcal{S}$. Hence for some $T \in D$, there exists a system of equations with no non-trivial solution in S_p for all $p \in T$. Since T is infinite (and hence non-empty) this is a contradiction. Hence Theorem 5 is proved, assuming the continuum hypothesis.

We now indicate how the assumption of the continuum hypothesis and the axiom of choice may be eliminated. It is easily shown that the statement of Theorem 5 is equivalent to an elementary number-theoretic statement; moreover the proof of this equivalence may be carried out in the set theory Σ described in Gödel [6]. Also the proof we have given that the continuum hypothesis and the axiom of choice imply the statement of Theorem 5 may be carried out in Σ. Now it is known (and follows easily from Gödel [6]) that a proof in Σ of an elementary number-theoretic statement which uses the continuum hypothesis and the axiom of choice may be transformed into a proof (in Σ) of the number-theoretic statement which does not use these assumptions. This completes the proof of Theorem 5.

The more direct way of eliminating the continuum hypothesis (but not the axiom of choice) by using Theorem 4′ was suggested to us by B. J. Birch. (See proof of Theorem 6.)

We turn now to a proof of the principle stated in the Introduction. For the convenience of the reader we first briefly describe the class of elementary statements for valued fields with a specified prime element π such that $\operatorname{ord}(\pi) = 1$, the smallest positive element of the valuation group. Let f, g be polynomials with fixed coefficients in the ring $Z[\pi]$. By an *atomic formula* we mean an expression of the form $\operatorname{ord}(f) = \operatorname{ord}(g)$, or $\operatorname{ord}(f) > \operatorname{ord}(g)$, or $\operatorname{ord}(f) = \infty$ (i.e. $f = 0$). An *elementary formula* is an expression constructed in a finite number of steps from atomic formulas by means of negation, conjunction, disjunction, implication, and quantifiers of the form "there exists an x such that" and "for all x." An *elementary statement* is an elementary formula involving no free variables, i.e. every

variable x in the formula is bound by a quantifier of the form "there exists an x such that" or "for all x." Given a valued field F with prime element π, we interpret an elementary statement for F by assuming that the variables in the statement range over the field; in addition if F has characteristic p we interpret any integers occurring in the statement modulo p.

As an example, let d and $n > d^2$ be fixed positive integers. Let f be the general form of degree d in the n variables $x_1, \cdots, x_n$ with (variable) coefficients $x_{n+1}, \cdots, x_t$ (so that $f \in Z[x_1, \cdots, x_t]$.) The following is an elementary statement which, by Theorem 5, is valid in all but a finite number of Q_p:

"For all $x_{n+1}, \cdots, x_t$ there exist $x_1, \cdots, x_n$ such that $f(x_1, \cdots, x_t) = 0$ and either $x_1 \neq 0$ or $x_2 \neq 0$ or $\cdots$ or $x_n \neq 0$."

The following property of ultraproducts accounts for their usefulness.

($\mathbf{II}$) An elementary statement is true for an ultraproduct if and only if it is true a.e. on the components. (For a proof see e.g. Kochen [9, Th. 5.1].)

It is this fact that enables us to reduce the proof of the following metamathematical principle to the algebraic statement that $\mathfrak{Q} \approx \mathfrak{S}$.

THEOREM 6. *For every elementary statement* Δ *there exists a finite set* C *such that, for all* $p \notin C$, Δ *is true in* Q_p *if and only if* Δ *is true in* S_p.

Proof. Let T be the set of primes p such that the elementary statement Δ holds in Q_p, and S the set of primes p such that Δ holds in S_p. We suppose, contrary to the theorem, that $(T - S) \cup (S - T)$ is infinite. Let D be a non-principal ultrafilter on the set P of primes with $(T - S) \cup (S - T)$ as an element. Then by using the property ($\mathbf{II}$) mentioned above, we have that Δ holds in one of $\prod_{p \in P} Q_p/D$ and $\prod_{p \in P} S_p/D$ but not the other. If we assume $2^{\aleph_0} = \aleph_1$ then this is impossible by Theorem 4.

To eliminate the continuum hypothesis we note that in the above proof we only need the fact that every elementary statement valid in $\mathfrak{Q}$ is valid in $\mathfrak{S}$ (in symbols $\mathfrak{Q} \equiv \mathfrak{S}$). Now by using the proof of the Löwenheim-Skolem theorem together with Theorem 4′ it is easily seen that there exist subfields $\mathfrak{Q}_\omega$, $\mathfrak{S}_\omega$ of $\mathfrak{Q}$, $\mathfrak{S}$ respectively such that $\mathfrak{Q}_\omega \equiv \mathfrak{Q}$, $\mathfrak{S}_\omega \equiv \mathfrak{S}$, and $\mathfrak{Q}_\omega$ and $\mathfrak{S}_\omega$ are analytically isomorphic. Hence $\mathfrak{Q} \equiv \mathfrak{S}$.

Clearly Theorems 1 and 5 are consequences of Theorem 6.

It is perhaps worth noting that Theorem 4 is a natural approach to establishing Theorem 6. We have seen that Theorem 4 implies Theorem 6; in fact the converse is also true. The precise sense in which this is true is contained in the following statement, which is a consequence of known results

in logic. Let $\{A_i \mid i \in I\}$ and $\{B_i \mid i \in I\}$ be two denumerably infinite families of algebraic systems for which $|A_i| \leqq 2^{\aleph_0}$ and $|B_i| \leqq 2^{\aleph_0}$, for all $i \in I$. If we assume the continuum hypothesis, then $\prod_{i \in I} A_i/D \cong \prod_{i \in I} B_i/D$, for all non-principal ultrafilters, is equivalent to the condition that for every elementary statement Δ, there exists a finite set C such that, for all $i \notin C$, Δ holds in A_i if and only if Δ holds in B_i.

CORNELL UNIVERSITY.

REFERENCES.

[1] J. Ax, "Zeroes of polynomials over finite fields," *American Journal of Mathematics*, vol. 86 (1964), pp. 255-261.

[2] ——— and S. Kochen, "Diophantine problems over local fields II," to appear in the *American Journal of Mathematics*.

[3] ——— and S. Kochen, "Diophantine problems over local fields III," to appear in the *Annals of Mathematics*.

[4] C. Chevalley, "Démonstration d'une hypothèse de M. Artin," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 11 (1936), pp. 73-75.

[5] B. Dwork, "On the zeta function of a hypersurface," *Institut des Hautes Études Scientifiques Publicationes Mathématiques*, No. 12 (1962), pp. 5-68.

[6] K. Gödel, *The consistency of the continuum hypothesis*, Annals of Mathematics Studies, Princeton University Press, 1940.

[7] H. Hasse, *Zahlentheorie*, Akademie-Verlag, Berlin 2nd ed. 1963.

[8] I. Kaplansky, "Maximal fields with valuations," *Duke Mathematical Journal*, vol. 9 (1942), pp. 313-321.

[9] S. Kochen, "Ultraproducts in the theory of models," *Annals of Mathematics* (2), vol. 74 (1961), pp. 221-261.

[10] W. Krull, "Allgemeine bewertungstheorie," *Journal für die Reine und Angewandte Mathematik*, vol. 167 (1932), pp. 160-196.

[11] S. Lang, "On quasi algebraic closure," *Annals of Mathematics* (2), vol. 55 (1952), pp. 373-390.

[12] ———, "Theorems and conjectures on diophantine equations," *Bulletin of the American Mathematical Society*, vol. 66 (1960), pp. 240-249.

[13] D. Lewis, "Cubic homogeneous polynomials over p-adic number fields," *Annals of Mathematics* (2), vol. 56 (1952), pp. 473-478.

[14] A. Ostrowski, "Untersuchungen zur arithmetischen Theorie der Körper," *Mathematische Zeitschrift*, vol. 39 (1934), pp. 269-404.

[15] K. Rychlik, "Zur Bewertungstheorie der algebraischen Körper," *Journal für die Reine und Angewandte Mathematik*, vol. 153 (1923), pp. 94-107.

[16] O. Schilling, *The Theory of Valuations*, American Mathematical Society Mathematical Surveys, No. IV, 1950.

[17] B. L. Van der Waerden, *Modern Algebra*, vol. 1, English Edition, Ungar, 1948.

[18] E. Weiss, *Algebraic Number Theory*, McGraw-Hill, New York, 1963.

DIOPHANTINE PROBLEMS OVER LOCAL FIELDS II. A COMPLETE SET OF AXIOMS FOR p-ADIC NUMBER THEORY.*

By James Ax and Simon Kochen.

---

0. Introduction. The notational conventions of [1] remain in force here. In that paper we gave a general method for handling diophantine assertions (or, more generally, elementary statements) which are true "semi-globally," i. e. which hold for all but a finite number of the fields Q_p of p-adic numbers. In this paper we study the class of elementary statements which hold in Q_p for a fixed prime p. We begin by axiomatizing Q_p. The axioms we adopt describe the following well-known properties of Q_p: Q_p is a valued field of characteristic 0 with valuation group a Z-group, R_p as residue class field, $\operatorname{ord}(p) = 1$, and satisfying the Hensel-Rychlik property. We call any valued field satisfying these properties a *formally p-adic field.* We shall see that these properties characterize Q_p in the sense that an elementary statement is valid in Q_p if and only if it is valid in every formally p-adic field. In fact, if an elementary statement is valid in Q_p, then there exists an elementary proof of it from the above axioms. (How the axioms themselves may put in an elementary form is indicated in Section 2). For example, if the conjecture of Artin stated in the first sentence of [1] is valid in Q_p (as is the case for $d = 3$ and $n > 9$), then it is valid in all formally p-adic fields, and (for each d and n) there is an elementary proof of the conjecture from the axioms. The situation here is entirely analogous to the one for the field of real numbers. The formally p-adic fields have the same relation to Q_p as the real-closed fields have to the field of reals.

It follows from the above-stated results for formally p-adic fields that there is an algorithm for deciding for every elementary statement whether or not it is valid in Q_p.[1] While this algorithm is useless as a tool for studying

---

Received August 24, 1964.

Revised April 15, 1965.

* While working on this paper the first author was partly supported by the U. S. Army Research Office (Durham) contract number DA-31-124-ARO-D-107 and National Science Foundation Grant GP-2243; the second author by National Science Foundation Grant GP-124. The authors thank S. Chase for discussing the homological aspect of this paper.

[1] The field Q_p is thus an example of a *decidable* field. The hitherto known examples of decidable fields are the finite fields, the algebraically closed fields, and the real closed

diophantine problems, it strongly hints at the possible existence of a "practical" algorithm. For example, in the analogous real field situation, the Sturm-Tarski-Seidenberg algorithm has some elements of practicality (see Jacobson [6, Chapter VI]).

We have written the above paragraphs with the non-logician in mind. In standard logical terminology the above results may be more succinctly stated as giving a complete, recursive axiomatization of the elementary theory of Q_p.

To prove this result we show, assuming the continuum hypothesis, that any two formally p-adic fields of cardinality $\leqq 2^{\aleph_0}$ have analytically isomorphic ultrapowers. This is proved by means of an isomorphism theorem (Theorem 1) similar to Theorem 2 of [1]. Two difficulties arise in adapting the methods of [1] to proving Theorem 1. The first is that we do not wish to postulate the existence of a cross-section. This necessitates proving the existence of such a cross-section (Lemma 7). The second difficulty lies in obtaining a result corresponding to the theorem of Kaplansky (Lemma 16 of [1]) for the present unequal characteristic case. Hypothesis A in Kaplansky [7] and [8] is too restrictive to apply in this situation. We introduce a new hypothesis which enables us to extend Kaplansky's theorem to cover, in particular, the present case of subfields of formally p-adic fields.

1. Formally p-adic fields.

Definition.[2] A *formally p-adic field* is a field V of characteristic 0, valued in a Z-group G, with residue class field R_p and satisfying the following properties:

(i) the Hensel-Rychlik property;

(ii) there is an algebraic closure $\tilde{V}$ of V such that $\tilde{V} = V\tilde{Q}$; and

(iii) $\operatorname{ord}(p) = 1$, the smallest positive element of G.

Q_p is a formally p-adic field. Using Lemma 10 of [1] and reasoning as in Lemma 14 of [1], we see that any algebraically closed subfield of a formally p-adic field is a formally p-adic field. In particular, the field of algebraic numbers of Q_p, $A_p = Q_p \cap \tilde{Q}$, is formally p-adic.

---

fields. Tarski has conjectured that these are the only decidable fields. (See J. Robinson, "The decision problem for fields," Symposium on the Theory of Models.) The formally p-adic fields are counter-examples to this conjecture. In [2], we shall give further examples of decidable fields.

[2] It turns out that axiom (ii) is actually a consequence of the other axioms. (See [2, Section 3].)

Throughout this section V is assumed to be a formally p-adic field with value group G. We also assume ord has been extended to avaluation of $\tilde{V}$. If F is a subfield of V, $\tilde{F}$ denotes the algebraic closure of F, taken in $\tilde{V}$.

LEMMA 1. *The following conditions on a valued field F are equivalent:*

(i) *F is formally p-adic and* $\operatorname{ord}(F)=Z$;

(ii) *F is formally p-adic and analytically isomorphic to a subfield of Q_p;*

(iii) *F is analytically isomorphic to a relatively algebraically closed subfield of Q_p.*

Proof. Condition (iii) implies (ii) by Lemma 10 of [1]. Also (ii) implies (i). Supposing (i) holds, we shall prove (iii). The usual Cauchy completion of F is a complete field with value group Z and residue class field R_p. Thus the completion of F is (analytically isomorphic to) Q_p. Let $\alpha \in Q_p$ be algebraic over F. We must show $\alpha \in F$. Let $J \in F[X]$ be the monic irreducible polynomial for α over F. By replacing α by $p^i\alpha$ with a sufficiently large integer i, we may assume that the coefficients of J are integral elements of F. Since Q is dense in Q_p, there exists $r \in Q \subseteq F$ such that $\operatorname{ord}(r) \geqq 0$ and $\operatorname{ord}(J(r)) > \operatorname{ord}(D(J))$. By condition (i) for the formally p-adic field F, J has a root in F. This proves Lemma 1.

Remark. There exists no non-trivial (algebraic) automorphism of a relatively algebraically closed subfield F of Q_p.

Proof. Since Q is dense in F, there exists no non-trival analytic automorphism of F. It therefore suffices to show that any (algebraic) automorphism ψ of F preserves the topology of F. Let U be the units of F, i. e. the set of $u \in F$ such that $\operatorname{ord}(u)=0$. Using the Hensel-Rychlik property for F, we find for $u \in F$ that $u \in U$ if and only if $u \in F^m$ for all positive integers m relatively prime to $p(p-1)$. This implies that U and hence the integers of F, $\mathfrak{D}=\bigcup_{i=0}^{\infty} p^iU$, are fixed under every automorphism of F. This verifies our remark.

LEMMA 2. *Let V be a formally p-adic field. Then there exists a unique analytic monomorphism $\psi\colon A_p \rightarrow V$. $\psi(A_p)$ is the field of (absolute) algebraic numbers of V.*

Proof. Let B be the field of (absolute) algebraic numbers of V. B satisfies (i) of Lemma 1. Thus there exists an analytic isomorphism $\Phi\colon B \rightarrow A$, where A is a relatively algebraically closed subfield of Q_p. Also

A is algebraic over Q. Thus $A = A_p$, and we may take $\psi = \Phi^{-1}$. As Q is dense in A_p and fixed by ψ, the uniqueness follows.

Notation. If m is a positive integer, let T'_m denote the set of all $(p^m - 1)$-th roots of unity, $T_m = T'_m \cup \{0\}$.

LEMMA 3. *Let* $\beta \in \tilde{Q}$, $n = [A_p(\beta) : A_p]$, $r = \operatorname{nord}(n)$,

$$e = (\operatorname{ord}(A_p(\beta)) : \operatorname{ord}(A_p)), \text{ and } f = [\overline{A_p(\beta)} : \bar{A}_p].$$

Then

(a) $n = ef$;

(b) *there exists* $\tau \in A_p(\beta) \cap T'_f$ *such that*

$$f = [A_p(\tau) : A_p] = [\overline{A_p(\tau)} : \bar{A}_p]; \text{ and}$$

(c) *there exist* $\pi \in A_p(\beta)$, $\tau_0 \in A_p(\beta) \cap T'_f$, *and* $\tau_1, \cdots, \tau_r \in A_p(\beta) \cap T_f$ *such that*

$$\pi^e = \left(\sum_{i=0}^{r} \tau_i p^i\right) p.$$

Proof. If we replace A_p by Q_p, (a) follows from Hasse [5, Chap. II, Section 14.5], (b) from [5, Chap. II, Section 10.3], and (c) from the Hensel-Rychlik theorem for $Q_p(\beta)$ (Rychlik [16]). We immediately deduce (b) and (c) for A_p using Lemma 10 of [1]. That condition (a) is true for A_p now follows from (b) and (c) and the inequality $ef \leqq n$. This proves the lemma.

LEMMA 4. *We may replace condition* (i) *in the definition of formally p-adic by*

$$\text{(i)}' \quad A_p = V \cap Q.$$

Proof. That (i)′ holds for any formally p-adic field, is merely a reformulation of Lemma 2.

Now assume that F satisfies the definition of a formally p-adic field with condition (i) replaced by (i)′. In order to show F satisfies (i), it suffices, by Lemma 12 of [1], to show that F satisfies the uniqueness property. We now do this in a fashion reminiscent of Lemma 11 of [1]. It suffices to show that if ord is extended to finite normal extension W of F and if α, $\alpha' \in W$ are F-conjugate, then $\operatorname{ord}(\alpha) = \operatorname{ord}(\alpha')$. By condition (ii), we may suppose (by taking an extension of W, if necessary) that $W = F(\beta)$, where $\beta \in \tilde{Q}$ and $A_p(\beta)/A_p$ is a finite normal extension. Let $n = [A_p(\beta) : A_p]$, $e = (\operatorname{ord}(A_p(\beta)) : \operatorname{ord}(A_p))$, and $f = [\overline{A_p(\beta)} : \bar{A}_p]$. We now select π, τ,

$\tau_0, \cdots, \tau_r$ as in Lemma 3. Let $\eta_{ij} = \pi^i\tau^j$, $i = 0, \cdots, e-1$, $j = 0, \cdots, f-1$. Then for all $f_{ij} \in F$, (since the η_{ij} form an integral basis for $F(\beta)/F$),

$$(*) \qquad \operatorname{ord}(\sum_{i,j} f_{ij}\eta_{ij}) = \min_{i,j}[\operatorname{ord}(f_{ij}) + \operatorname{ord}(\eta_{ij})].$$

Also $\operatorname{ord}(\eta_{ij}) = i/e$. This last fact is seen from (c) of Lemma 3, since there is a unique extension of ord to $Q_p(\tau)$. Thus

$$(**) \qquad \operatorname{ord}(\sum_{i,j} f_{ij}\eta_{ij}) = \min_{i,j}[\operatorname{ord}(f_{ij}) + i/e].$$

Now the conjugates π', τ', $\tau'_0, \cdots, \tau'_r$ of $\pi, \tau, \tau_0, \cdots, \tau_r$ also satisfy Lemma 3, so that (**) holds with $\eta'_{ij} = \pi'^i\tau'^j$. This completes the proof of the lemma.

As a corollary to the proof of Lemma 4 we obtain the following fact.

LEMMA 5. *V has the uniqueness property.*

PROPOSITION 1. *Let F be a relatively algebraically closed subfield of V. If $\alpha \in \bar{F}$ and $y \in V$, then there exists $\phi \in F$ such that* $\operatorname{ord}(y - \phi) \geqq \operatorname{ord}(y - \alpha)$.

Proof. By condition (ii) for the formally p-adic field V, there exists $\beta \in \bar{Q}$ such that $\alpha \in L = V(\beta)$. As in the proof of Lemma 4, there is an integral basis η_i, $i = 1, \cdots, n$, with $\eta_1 = 1$, for L/V consisting of elements of $A_p(\beta)$. Since $\alpha \in L = VF(\beta)$, it follows from Lemma 10 of [1] that $\alpha \in F(\beta)$. As the η_i form a basis for L/V, there exist $f_i \in F$ such that $\alpha = \sum_i f_i\eta_i$. Defining

$$d_i = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{otherwise} \end{cases}$$

we have

$$\alpha - y = \sum_i (f_i - d_i y)\eta_i.$$

Now

$$\begin{aligned} \operatorname{ord}(\alpha - y) &= \min_i(\operatorname{ord}(f_i - d_i y)\eta_i) \\ &\leqq \operatorname{ord}(f_1 - d_1 y) = \operatorname{ord}(f_1 - y). \end{aligned}$$

Thus $\phi = f_1$ is an element of F with the required property.

Definition. A subfield F of V is *V-maximally complete* if there is no immediate extension of F in V.

Remark. If F is V-maximally complete and $\operatorname{ord}(F)$ is pure in G, then F is relatively algebraically closed in V.

PROPOSITION 2. *Assume V is ω-pseudo-complete. If F is V-maximally complete and* $\operatorname{ord}(F) = H$ *is a pure, countable subgroup of G, then F is maximally complete.*

Proof. The proof of Proposition 2 of [1], with the notation there understood in its (slightly different) present sense, is a proof of the present proposition.

LEMMA 6. *Assume V is ω-pseudo-complete. Then there exists an analytic monomorphism of Q_p into V.*

Proof. Let Z be the subgroup $\langle 1 \rangle$ of G generated by 1. Let F be a subfield of Q_p which is maximal with respect to the property: $\mathrm{ord}(F) \subseteq Z \subseteq G$. Then F is V-maximally complete. Also Z is a pure, countable subgroup of G. It follows from Proposition 2 that F is maximally complete. Hence F is complete. As $\mathrm{ord}(F) = Z$ and $\bar{F} = R_p$, we have that F is analytically isomorphic to Q_p. This proves the lemma.

LEMMA 7. *Assume V is ω-pseudo-complete. Then there exists a monomorphism π of the group $G = \mathrm{ord}(V)$ into V satisfying the following conditions:*

(i) $\mathrm{ord}(\pi(\gamma)) = \gamma$ *for all* $\gamma \in G$.

(ii) $\pi(1) = p$.

In other words, V has a "normalized" cross-section.

Proof Let U be the group of units of V, and let i be the injection of U into V. To prove (i), we must show that the short exact sequence

$$(*) \qquad 0 \to U \xrightarrow{\ i\ } V \xrightarrow{\ \mathrm{ord}\ } G \to 0$$

splits.

It suffices to show $\mathrm{Ext}_Z{}^1(G, U) = 0$. By Lemma 6, we may regard Q_p as a subfield of V. Let U_p be the group of units of Q_p. Let U' be the subgroup of U consisting of all $u \in U$ such that $\mathrm{ord}(u-1) > n$ for all $n \in Z \subseteq G$. Given $u \in U$, define $\tau(u) = \sum_{i=0}^{\infty} a_i p^i \in U_p$ (where the $a_i \in Z$ and $0 < a_0 < p$, $0 \leqq a_i < p$, for $i = 1, 2, \cdots$), by means of the equalities,

$$\bar{u} = \bar{a}_0$$

and

$$\overline{(u - \sum_{i=0}^{n-1} a_i p^i)/p^n} = \bar{a}_n, \qquad \text{if } n > 0.$$

We see that

$$(**) \qquad U = U_p U' \text{ (internal, direct)};$$

and for $u \in U$

$$u \to (\tau(u), u/\tau(u)) \in U_p \oplus U'$$

is an isomorphism of U onto the (external) direct sum $U_p \oplus U'$.

Now $U_p \sim T_p \oplus Z_p$, where T_p is a finite cyclic group (see, for example, [5, Chap. II, Section 15.7]). Thus

$$(**) \qquad U \approx T_p \oplus Z_p \oplus U'.$$

Hence,

$$\mathrm{Ext}_Z{}^1(G, U) \approx \mathrm{Ext}_Z{}^1(G, T_p) \oplus \mathrm{Ext}_Z{}^1(G, Z_p) \oplus \mathrm{Ext}_Z{}^1(G, U').$$

For any abelian group A we denote the group of all homomorphisms of A into the multiplicative group of all complex numbers of absolute value 1 by $D(A)$. By Cartan and Eilenberg [3, Chap. VII, Prop. 6.1], we have for all abelian groups A and B (dropping the subscript "Z"),

$$\mathrm{Ext}^1(A, D(B)) \approx D(\mathrm{Tor}_1(A, B)).$$

Now $T_p \approx D(T_p)$ and $Z_p \approx D(Q_p/Z_p)$. As G is torsion-free, $\mathrm{Tor}_1(G, B) = 0$ for all abelian groups B [3, Chap. VII, Prop. 4.2]. It follows that

$$\mathrm{Ext}^1(G, T_p) = \mathrm{Ext}^1(G, Z_p) = 0.$$

These equalities also follow from some results of Nunke [12, Theorem 7.4].

To establish (i) it now suffices to prove U' is divisible, for then by [3, Chap. VI, Cor. 2.2a] $\mathrm{Ext}^1(G, U') = 0$. Thus, let n be a positive integer and let $u \in U'$. Then

$$\mathrm{ord}(1^n - u) > n\,\mathrm{ord}(n),$$

and so by the Hensel-Rychlik property for V, there exists a $v \in V$ such that $v^n = u$. Evidently $v \in U$. By (**) we may assume $v \in U'$. This shows U' is divisible and completes the proof of (i).

Thus there exists a homomorphism $\pi_0 \colon G \to V$ such that $\mathrm{ord}(\pi_0(\gamma)) = \gamma$ for all $\gamma \in G$. Now let a homomorphism $h_0 \colon Z \to U$ be defined by requiring $h_0(1) = p/\pi_0(1) \in U$.

We claim there exists an extension h of h_0 to a homomorphism of G into U. Indeed, the exact sequence

$$0 \to Z \xrightarrow{\ j\ } G \xrightarrow{\ \beta\ } G/Z \to 0$$

induces an exact sequence

$$\operatorname{Hom}(G,U) \xrightarrow{\delta} \operatorname{Hom}(Z,U) \to \operatorname{Ext}^1(G/Z,U)$$

where $\delta = \operatorname{Hom}(j,\lambda)$ with λ the identity map of U. Now the above proof that $\operatorname{Ext}^1(G,U) = 0$ used only the fact that G is torsion free and hence shows that $\operatorname{Ext}^1(G/Z,U) = 0$, because Z is a pure subgroup of G. Thus δ is onto. This proves that the desired extension h of h_0 exists. We now define $\pi\colon G \to V$ by $\pi(\gamma) = h(\gamma)\pi_0(\gamma)$ for all $\gamma \in G$. Since $h(\gamma) \in U$ for all $\gamma \in G$, condition (i) holds for π. As $\pi(1) = h(1)\pi_0(1) = p$, (ii) holds for π. This proves the lemma.

Definition. If E is a subfield of V, we denote by $[E \mid V]$ some fixed subfield L of V, maximal with respect to the property: $\operatorname{ord}(L) \subseteq (\operatorname{ord}(E) \mid G)$.

Remark. $[E \mid V]$ is relatively algebraically closed in V.

We now need to know that subfields of formally p-adic fields have ***unique*** immediate maximal completions. Unfortunately, neither Lemma 16 of [1], nor its generalization—Kaplansky's results [7, Th. 5]—cover the present situation, because of the restrictiveness of Kaplansky's hypothesis A. To overcome this difficulty we introduce a hypothesis which is satisfied in the present situation; we are thereby enabled to obtain the desired uniqueness statement (Lemma 11), by copying the proof in [7].

We shall refer to the following condition on a field K of characteristic 0, valued in a group Γ as *hypothesis B*.

For all $j \in Z \subseteq K$ there are only finitely many $\gamma \in \Gamma$ such that $\operatorname{ord}(j) \geqq \gamma \geqq 0$.

LEMMA 8. *Suppose K satisfies hypothesis B. Let $a_\rho \in K$ form a pseudo-Cauchy sequence. Suppose that $q(X) \in K[X]$ is a polynomial of smallest degree n such that* $\operatorname{ord}(a_\rho)$ *eventually increases (i.e., $\{a_\rho\}$ is of* algebraic type *with* $q(X)$ *as a* minimal polynomial). *Let $\beta_i \in \Gamma$ be the eventual value of* $\operatorname{ord}(q^{(i)}(a_\rho)/i!)$. *We may conclude that if $j \in Z$, $j > 1$ and if $\gamma_\rho = \operatorname{ord}(a_{\rho+1} - a_\rho)$, then $\beta_1 + \gamma_\rho < \beta_j + j\gamma_\rho$ eventually.*

Proof. We form the Taylor expansion of $q^{(1)}(X)$ about a_ρ and evaluate at $a_{\rho+1}$:

$$\begin{aligned} q^{(1)}(a_{\rho+1}) - q^{(1)}(a_\rho) &= \sum_{i=1}^{n} (q^{(i+1)}(a_\rho)/i!)(a_{\rho+1} - a_\rho)^i \\ &= \sum_{i=1}^{n} (i+1)(q^{(i+1)}(a_\rho)/(i+1)!)(a_{\rho+1} - a_\rho)^i. \end{aligned}$$

By [7, Lemma 4], eventually there will be among the summands of the right member, precisely one of least ord. The ord of this term must then equal the ord of the left member, which is not less than β_1. It follows that

$$\beta_1 \leqq \operatorname{ord}(j(q^{(j)}(a_\rho)/j!)(a_{\rho+1}-a_\rho)^{j-1}).$$

Thus,

$$\beta_1+\gamma_\rho \leqq \operatorname{ord}(j)+\beta_j+j\gamma_\rho \text{ eventually.}$$

Now if this relation holds for $\rho \geqq \rho_0$ and if M is the (finite) number of $\gamma \in \Gamma$ such that $\operatorname{ord}(j) \geqq \gamma \geqq 0$, then $\beta_1+\gamma_\rho < \beta_j+j\gamma_\rho$, for $\rho \geqq \rho_0+M$, since the γ_ρ are strictly increasing.

LEMMA 9. *Assume the notation and assumptions of Lemma 8. In addition, let N be an immediate maximal completion of K. If $y \in N$ is a pseudo-limit of $\{a_\rho\}$, then*

(i) $\operatorname{ord}(q(y)) > \beta_1+\gamma_\rho$, *for all* ρ, *and*

(ii) $\operatorname{ord}(q^{(i)}(y)/i!) = \beta_i$ *for* $i=1,2,\cdots,n$.

Proof. We have

$$(*) \qquad q(x)-q(a_\rho)=\sum_{i=1}^{n}(q^{(i)}(a_\rho)/i!)(x-a_\rho)^i, \text{ for all } x \in N.$$

Let $\sigma > \rho$ and let $x=a_\sigma$. It follows from Lemma 8 that

$$\begin{aligned}\operatorname{ord}(q(a_\sigma)-q(a_\rho)) &= \operatorname{ord}(q^{(1)}(a_\rho)(a_\sigma-a_\rho))\\ &= \beta_1+\gamma_\rho \quad \text{for } \rho \geqq \rho_0, \text{ say.}\end{aligned}$$

Varying σ, while fixing $\rho \geqq \rho_0$, we know $\operatorname{ord}(q(a_\sigma))$ increases eventually. This implies that

$$(**) \qquad \operatorname{ord}(q(a_\rho)) = \beta_1+\gamma_\rho \quad \text{for } \rho \geqq \rho_0.$$

Now let $x=y$ in (*). Since y is a pseudo-limit of $\{a_\rho\}$, $\operatorname{ord}(y-a_\rho) = \gamma_\rho$, for all ρ. It follows from (*) that $\operatorname{ord}(q(y)-q(a_\rho))=\beta_1+\gamma_\rho$ for $\rho \geqq \rho_0$. Statement (**) now implies that $\operatorname{ord}(q(y)) \geqq \beta_1+\gamma_\rho$ for $\rho \geqq \rho_0$. Statement (i) is an immediate consequence of this relation.

To prove (ii) we use the Taylor expansion:

$$\begin{aligned}q^{(i)}(y)/i! - q^{(i)}(a_\rho)/i! &= (1/i!)\sum_{j=i+1}^{n}(q^{(j)}(a_\rho)/(j-i)!)(y-a_\rho)^{j-i}\\ &= \sum_{j=i+1}^{n}\binom{j}{i}(q^{(j)}(a_\rho)/j!)(y-a_\rho)^{j-i}.\end{aligned}$$

By [7, Lemma 4], the right member increases eventually. Hence

$$\operatorname{ord}(q^{(i)}(y)/i!) = \operatorname{ord}(q^{(i)}(a_\rho)/i!) = \beta_1$$

eventually. This proves Lemma 9.

LEMMA 10. *Assume the notation and assumptions of Lemma 9. Let $t \in N$ be a pseudo-limit of $\{a_\rho\}$ with* $\operatorname{ord}(q(t)) = \alpha \in \Gamma$. *Then there exists $t^* \in N$ satisfying the following conditions:*

(i) *t^* is a pseudo-limit of $\{a_\rho\}$.*

(ii) $\operatorname{ord}(t^* - t) = \alpha - \beta_1$.

(iii) $\operatorname{ord}(q(t^*)) > \alpha$.

Proof. By Lemma 9,

$$\alpha = \operatorname{ord}(q(t)) > \beta_1 + \gamma_\rho, \quad \text{for all } \rho.$$

Hence

$$(*) \qquad \alpha - \beta_1 > \gamma_\rho, \quad \text{for all } \rho.$$

Let $k \in N$ be such that $\operatorname{ord}(k) = \alpha - \beta_1$. For Z a variable over N, we have

$$(**) \qquad r(Z) = q(t + kZ)/q(t) = \sum_{j=0}^{n} k^j Z^j q^{(j)}(t)/[(j!)q(t)].$$

In $r(Z)$ the coefficient of Z^j has ordinal $\delta(j) = j(\alpha - \beta_1) + \beta_j - \alpha$ for $j \geqq 1$. Also, $\delta(1) = 0$. From $(*)$ and Lemma 6, we have, for $j > 1$,

$$\delta(j) = (j-1)(\alpha - \beta_1) + \beta_j - \beta_1 > (j-1)\gamma_\rho + \beta_j - \beta_1 > 0 \text{ eventually.}$$

This shows that $\bar{r}(Z) \in \bar{N}[Z]$ has degree precisely 1. Let $\zeta \in N$ be a root of $\bar{r}$. Let $z \in N$ be such that $\bar{z} = \zeta$. We define $t^* = t + kz$. It follows from $(*)$ and the definitions of k and z that (i) and (ii) hold. Statement (iii) follows from $(**)$ and our choice of z. This proves Lemma 10.

LEMMA 11. *If K satisfies hypothesis B, then there is a* unique *immediate maximally complete extension of K (up to analytic equivalence).*

Proof. The existence was proved by Krull [11, Th. 24]. According to question $(*)$ of [7, Section 3, p. 309] and the statement following $(*)$, the uniqueness will follow if we can show, in the notation of Lemma 10, that N contains a pseudo-limit of $\{a_\rho\}$ which is also a root of $q(X)$.

Assume N does not contain a pseudo-limit of $\{a_\rho\}$ which is also a root of $q(X)$; we will show this leads to a contradiction. Let λ be an ordinal

of cardinality exceeding $|\Gamma|$. We are going to pick $t_\mu \in N$, for $\mu < \lambda$ such that

(1) each t_μ is a pseudo-limit of $\{a_\rho\}$;

(2) if $\operatorname{ord}(q(t_\mu)) = \alpha_\mu$, then $\alpha_\mu < \alpha_\nu$, for $\mu < \nu < \lambda$;

(3) $\operatorname{ord}(t_\nu - t_\mu) = \alpha_\mu - \beta_1$, for $\mu < \nu < \lambda$.

This will give the desired contradiction.

We let t_0 be any pseudo-limit of $\{a_\rho\}$. If $\sigma < \lambda$ and t_σ has already been chosen, we apply Lemma 10 with $t = t_\sigma$ and set $t_{\sigma+1} = t^*$. If $\sigma < \lambda$ is a limit ordinal, and if t_μ has already been chosen for all $\mu < \sigma$, then the sequence $\{t_\mu\}_{\mu<\sigma}$ is pseudo-Cauchy. We let t_σ be any pseudo-limit of $\{t_\mu\}$ in N. As in [7, pp. 313-314], these choices satisfy the requirements (1), (2), and (3). This proves Lemma 11.

PROPOSITION 3. *Let V and V' be ω-pseudo-complete formally p-adic fields with common value group G. Let F (resp.: F') be V-maximally complete (resp.: V'-maximally complete) and suppose* $\operatorname{ord}(F)$ *is a countable, pure subgroup of G. Let $\psi\colon F \to F'$ be an analytic isomorphism. If $x \in V - F$, then there exists an extension ψ_0 of ψ to an analytic isomorphism of $[F(x) \mid V]$ onto some subfield F_1' of V'. Also,*

$$\operatorname{ord}([F(x) \mid V]) = (\operatorname{ord}(F(x)) \mid G) = (\operatorname{ord}(F_1') \mid G).$$

Proof. The Remark preceding Proposition 2 shows that F (resp.: F') is a formally p-adic field. The V-maximal completeness of F implies that $\operatorname{ord}(F(x)) \supsetneq \operatorname{ord}(F)$. Thus there exists $A \in F[X]$ such that $\operatorname{ord}(A(x)) \in G - \operatorname{ord}(F)$. As $\operatorname{ord}(F)$ is a pure subgroup of G, this relation is equivalent to $\operatorname{ord}(A(x)) \in G - (\operatorname{ord}(F) \mid G)$, which in turn implies $\operatorname{ord}(A(x)) \notin \operatorname{ord}(F)$. By factoring A over $\tilde{F}$, we see there exists an $\alpha \in \tilde{F}$ such that $\operatorname{ord}(x-\alpha) \notin \operatorname{ord}(\tilde{F})$. By Proposition 1, there is an $f \in F$ such that $\operatorname{ord}(x-f) \geqq \operatorname{ord}(x-\alpha)$. If $\operatorname{ord}(x-f) > \operatorname{ord}(x-\alpha)$, then

$$\operatorname{ord}(f-\alpha) = \operatorname{ord}(x-\alpha) \notin \operatorname{ord}(\tilde{F}),$$

a contradiction. Thus

$$\operatorname{ord}(x-f) = \operatorname{ord}(x-\alpha) \in G - (\operatorname{ord}(F) \mid G).$$

Replacing x by $x-f$, we may assume $\gamma = \operatorname{ord}(x) \in G - (\operatorname{ord}(F) \mid G) = G - \operatorname{ord}(F)$. By Lemma 7, there exists a cross-section π (resp.: π') into V (resp.: V') such that $\pi(1) = p \in V$ (resp.: $\pi'(1) = p \in V'$).

By Proposition 2, F is maximally complete. By Lemma 6 and Lemma 15

of [1] we may regard Q_p as a subfield of F. Since $\psi(Q_p) \subseteq F'$, we may regard Q_p as a common subfield of F and F', fixed by ψ.

Since $\operatorname{ord}(x\pi(-\gamma)) = 0$ there exists (as in the proof of (**) in Lemma 7) $c \in U_p \subset Q_p{}^*$ such that $\operatorname{ord}(cx\pi(-\gamma) - 1) > m$ for all positive integers m. Replacing x by cx, we may assume $\operatorname{ord}(x\pi(-\gamma) - 1) > m$ for all positive integers m. We set $x' = \pi'(-\gamma)$. Then there exists a unique analytic isomorphism $\psi_1: F(x) \to F'(x')$ which extends ψ and such that $\psi_1(x) = x'$. Indeed, since $\gamma \notin (\operatorname{ord}(F) \mid G)$, x (resp.: x') is transcendental over F (resp.: F'). Thus ψ_1 is an algebraic isomorphism. If $A(X) = \sum a_i X^i \in F[X]$, we have

$$\begin{aligned}\operatorname{ord}(A(x)) &= \min_i(\operatorname{ord}(a_i) + i\gamma)\\ &= \min_i(\operatorname{ord}(\psi(a_i) + i\gamma) = \operatorname{ord}(\sum \psi(a_i)x'^i)\\ &= \operatorname{ord}\psi_1(A(x)),\end{aligned}$$

so ψ_1 is an analytic isomorphism.

Since G is a Z-group, for all positive integers n, there exist unique $\gamma(n) \in G$ and $r(n) \in Z$ such that

$$\gamma = n\gamma(n) + r(n), \qquad 0 \leqq r(n) < \mathrm{n}.$$

We define $y_n = x\pi(-r(n)) = xp^{-r(n)} \in F(x)$ (resp.: $y_n' = \psi_1(y_n) = x'p^{-r(n)} = \pi'(n\gamma(n)) \in F'(x')$). Thus, $\operatorname{ord}(y_n\pi(-\gamma(n))^n - 1) = \operatorname{ord}(x\pi(-\gamma) - 1) > n\operatorname{ord}(n)$. By the Hensel-Rychlik property, $y_n\pi(-\gamma(n))^n$ has an n-th root in V. Thus y_n has an n-th root z_n in V and hence in $[F(x) \mid V]$. The field $F(z_n) \supseteq F(x)$ is independent of which n-th root of y_n is chosen since all roots of unity in V are already in $A_p \subseteq F$. We let $z_n' = \pi'(\gamma(n))$. Thus z_n' is an n-th root of y_n'. Hence, $z_n' \in [F'(x') \mid V']$.

We assert that, for each positive integer n, there exists an extension of ψ_1 to an analytic isomorphism $\sigma: F(z_n) \to F'(z_n')$ defined by requiring $\sigma(z_n) = z_n'$. The map σ is an algebraic isomorphism because $z_n{}^n = y_n$, $z_n'^n = y_n' = \psi_1(y_n)$, and, as we shall see, $X^n - y_n$ is irreducible over $F(x)$. To see that $X^n - y_n$ is irreducible over $F(x)$, it suffices to show that $(\operatorname{ord}(F(z_n)) : \operatorname{ord}(F(x))) \geqq n$, i.e. $0 < j < n$ implies $j\gamma(n) \notin \operatorname{ord}(F(x))$. This will also show that σ is analytic. Assume $m \in Z$ and $\delta \in \operatorname{ord}(F)$ are such that

$$j\gamma(n) = \delta + m\gamma \in \operatorname{ord}(F(x)).$$

Then $j\gamma(n) = \delta + m(n\gamma(n) + r(n)) = mn\gamma(n) + \delta + mr(n)$. Thus

$$(j - mn)\gamma(n) = \delta + mr(n) \in \operatorname{ord}(F).$$

It follows that

$$\begin{aligned}(j-mi)\gamma &= (j-mn)(n\gamma(n)+r(n))\\ &= n(j-mn)\gamma(n)+(j-mn)r(n)\in \operatorname{ord}(F).\end{aligned}$$

Since $0<j<n$, we have $\gamma\in(\operatorname{ord}(F)\mid G)$, a contradiction. This proves our assertion that the set Σ_n of analytic isomorphisms $\sigma\colon F(z_n)\to F'(z_n')$ extending ψ_1 is non-empty. Thus $1\leqq|\Sigma_n|\leqq$ n.

We now assert there exists an extension of ψ_1 to an analytic isomorphism $\psi_2\colon F(z_n\colon n=1,2,\cdots)\to F'(z_n'\colon n=1,2,\cdots)$. We give each Σ_n the discrete topology and $\Sigma=\prod_{n=1}^{\infty}\Sigma_n$, the product topology. By Tychonoff's Theorem, Σ is compact. For each natural number n, let $T(n)$ be the (closed) set of $t\in\Sigma$ such that $t(m)=t(n)\mid F(z_m)$ if $F(z_m)\subseteq F(z_n)$. Suppose $d\mid n$, say $n=de$. Then $(z_n{}^e)^d=y_n$. Let $a,r\in Z$ be such that

$$r(n)=ad+r. \qquad\qquad 0\leqq r<d.$$

Thus

$$\gamma=n\gamma(n)+r(n)=n\gamma(n)+ad+r=d(e\gamma(n)+a)+r.$$

This implies $r=r(d)=r(n)-ad$. It follows that $(z_n{}^e)^d=x\pi(-r(n))=x\pi(-r(d)-ad)$. Hence $(z_n{}^e\pi(a))^d=x\pi(-r(d))=y_d$. Thus $F(z_n)\supseteq F(z_d)$. This implies that $T(n)\cap T(n')\supseteq T(nn')$, for all positive integers n and n', so that the $T(n)$ have the finite intersection property. Thus $\bigcap_{n=1}^{\infty}T(n)\neq\phi$. Any $t\in\bigcap_{n=1}^{\infty}T(n)$ defines an extension of ψ_1 to an analytic isomorphism $\psi_2\colon F(z_n\colon n=1,2,\cdots)\to F'(z_n'\colon n=1,2,\cdots)$ by requiring that $\psi_2\mid F(z_n)=t(n)$.

We next claim that $\operatorname{ord}(F(z_n\colon n=1,2,\cdots))=(\operatorname{ord}(F(x))\mid G)$. Let $\delta\in(\operatorname{ord}(F(x))\mid G)$ and let m be a positive integer such that $m\delta\in\operatorname{ord}(F(x))$, say $m\delta=h+n\gamma$ with $h\in\operatorname{ord}(F)$, and $n\in Z$. Let $h_1\in G$ and $m_1\in Z$ be such that

$$h=mh_1+m_1, \qquad\qquad 0\leqq m_1<m.$$

Since $\operatorname{ord}(F)$ is a pure subgroup of G and $h-m_1\in\operatorname{ord}(F)$, we see $h_1\in\operatorname{ord}(F)$. We have

$$\begin{aligned}m\delta &= mh_1+m_1+n\gamma\\ &= mh_1+m_1+n(m\gamma(m)+r(m))=m(h_1+n\gamma(m))+m_1+nr(m).\end{aligned}$$

It follows that $m\mid(m_1+nr(m))$ and

$$\delta=h_1+n\gamma(m)+(m_1+nr(m))/m.$$

Thus $\delta \in \operatorname{ord}(F(z_n\colon n=1,2,\cdots))$. This shows $\operatorname{ord}(F(z_n\colon n=1,2,\cdots)$ $\supseteq (\operatorname{ord}(F(x)) \mid G)$. The reverse inclusion is a direct consequence of the definitions. This establishes our claim.

We may now conclude that $\operatorname{ord}([F(x) \mid V]) = (\operatorname{ord}(F(x) \mid G)$ and that $[F(x) \mid V]$ (resp.: $[F'(x') \mid V'])$ is an immediate extension of $F(z_n\colon n=1,2,\cdots)$ (resp.: $F'(z_n'\colon n=1,2,\cdots)$. As $\operatorname{ord}(F)$ is countable, so is $(\operatorname{ord}(F(x)) \mid G)$. It follows from Proposition 2 that $[F(x) \mid V]$ (resp.: $[F'(x') \mid V])$ is maximally complete. By Lemma 11, there exists an extension ψ_0 of ψ_2,

$$\psi_0\colon [F(x) \mid V] \to [F'(x') \mid V'] = F_1'$$

which is an analytic isomorphism.

This completes the proof of Proposition 3.

THEOREM 1. *Let V and V' be two ω-pseudo-complete formally p-adic fields of cardinality $\aleph_1$ with the same value group G of cardinality $\aleph_1$. Then there exists an analytic isomorphism*

$$\psi\colon V \to V'.$$

Moreover, if $F \subset V$ (resp.: $F' \subset V'$) *is a formally p-adic field such that* $\operatorname{ord}(F)$ *is countable, and if $\phi\colon F \to F'$ is an analytic isomorphism, then ψ may be taken to be an extension of ϕ.*

Proof. By Lemma 2, it suffices to prove the second assertion. The proof is now entirely similar to the proof of Theorem 2 of [1]. In fact, in the proof of Theorem 2 of [1], if we systematically substitute F for $C(\rho)$ (resp.: F' for $C(\rho')$) and if we interpret the reference to Proposition 3 there to refer to Proposition 3 of this paper, then we obtain a proof of the present theorem.

We shall now identify the formally p-adic fields described in the statement of Theorem 1. In a natural way we regard Z as a subgroup of G. Let H denote the quotient group. H inherits an ordering since Z is an *isolated* subgroup of G. (See [17, p. 5].) For each $h \in H$, let $\alpha(h)$ be a fixed preimage of h in G. Define a 2-cocycle β of H in Z by $\beta(h,h') = \alpha(h+h') - \alpha(h) - \alpha(h')$ for all $h, h' \in H$. Then G is order-isomorphic to the group of couples $(z,h) \in Z \times H$ with addition defined by $(z,h) + (z',h') = (z+z'+\beta(h,h'), h+h')$ and with lexicographic ordering:

$$(z,h) \leqq (z',h') \Longleftrightarrow h \leqq h' \text{ or } h=h' \text{ and } z \leqq z'.$$

Definition. Let δ be an infinite cardinal. $Q_p(t^H;\beta)_\delta$ denotes the valued

field of formal sums $a = \sum_{h \in S} c_h t^h$, where S is an arbitrary well-ordered subset of H of cardinality at most δ and $c_h \in Q_p$ for all $h \in S$. Here multiplication is defined formally except that it is required that

$$c_h t^h c'_h t^{h'} = (c_h c'_h p^{\beta(h,h')}) t^{h+h'}.$$

We define the valuation on $Q_p(t^H;\beta)_\delta$ by $\operatorname{ord}(a) = (\operatorname{ord}_p(c_{h_0}), h_0)$ where $\operatorname{ord}_p\colon Q_p \to Z \cup \{\infty\}$ is the natural valuation and h_0 is the smallest element of S such that $c_{h_0} \neq 0$.

THEOREM 2. *If V satisfies the hypothesis of Theorem 1, then V is analytically isomorphic to $Q_p(t^H;\beta)_{\aleph_0}$, assuming the continuum hypothesis.*

The proof is similar to the proof of Theorem 3 of [1].

2. The completeness of the axioms. Our first task is to reformulate the axioms for formally p-adic fields by means of (an infinite set of) elementary statements. Such a formulation for the case of all valued fields is straightforward (see A. Robinson [14, Section 3.4]). The remaining axioms, for formally p-adic fields, with the exception of axiom (ii), can clearly be put in elementary form.

We shall need to refine condition (ii) for formally p-adic fields.

LEMMA 12. *Let V be a formally p-adic field. Let $V \subseteq W \subset \bar{V}$, $[W:V] = n < \infty$. Then there exists $\alpha \in \bar{Q}$ such that $W = V(\alpha)$. Moreover, $[A_p(\alpha) : A_p] = n$.*

Proof. By Lemma 10 of [1], the second assertion is a consequence of the first. By condition (ii) for the formally p-adic field V, there exists $\beta \in Q$ such that $W \subseteq V(\beta)$. We may assume that $A_p(\beta)/A_p$ is a normal extension with Galois group M. $V(\beta)/V$ is a normal extension with Galois group N monomorphic to M under the natural "restriction" monomorphism ρ. By Lemma 10 of [1], $[V(\beta):V] = [A_p(\beta):A_p]$. Hence $\rho\colon N \to M$ is an isomorphism. It follows now that the lattice of sub-extensions of $V(\beta)/V$ is isomorphic to the lattice of sub-extensions of $A_p(\beta)/A_p$. If N_1 is the subgroup of N fixing W, then any primitive element α for the fixed field of $\rho(N_1)$ satisfies the requirement of the lemma.

LEMMA 13. *Let V be a formally p-adic field. Suppose $V \subseteq W \subset \tilde{V}$ with $[W:V] = n < \infty$. Let $r = \operatorname{nord}(n)$. Then there exist*

$$\pi, \sigma, \tau_0, \tau_1, \cdots, \tau_r \in Q$$

such that:

(1) *σ is a primitive $(p^n - 1)$-th root of unity;*

(2) $\tau_0 \in T_n'$, $\tau_i \in T_n$, $i = 1, \cdots, r$;

(3) $\pi^n = (\sum_{i=0}^{r} \tau_i p^i) p$; *and*

(4) $W \subseteq V(\sigma + \pi)$.

Proof. We first show that $V(\sigma+\pi) = V(\sigma,\pi)$. Since, by Lemma 5, V has the uniqueness property, $V(\sigma+\pi)$ also has the uniqueness property. Hence, by Lemma 12 of [1], $V(\sigma+\pi)$ has the Hensel-Rychlik property. Now, $\bar{\sigma} = \sigma + \pi \in \overline{V(\sigma+\pi)}$, so that by the Hensel-Rychlik property for $V(\sigma+\pi)$, $\sigma \in V(\sigma+\pi)$. Thus, $\pi \in V(\sigma+\pi)$, and $V(\sigma+\pi) = V(\sigma,\pi)$. This lemma now follows from Lemma 3 and Lemma 12 since when $n = ef$, a $(p^f - 1)$-th root of unity is a $(p^n - 1)$-th root of unity and an e-th root of $(\sum_{i=0}^{r} \tau_i p^i) p$ is the f-th power of some n-th root.

Let $[W:V] = n$. It follows from Lemma 13 that there is a set $L(n)$ of absolute alegrabic integers $\sigma + \pi$ such that:

(a) $\lambda = \lambda(n) = |L(n)| \leqq n\phi(p^n - 1)(p^n - 1)p^{n^2 \operatorname{ord}(n)}$ where ϕ is the Euler function; and

(b) there exists $\alpha \in L(n)$ such that $W \subseteq V(\alpha)$.

The monic irreducible polynomials for the $\alpha \in L(n)$ over V have rational integral coefficients and can be explicitly computed. Let $M(n)$ denote the resulting set of $\lambda(n)$ polynomials. We may now reformulate condition (ii) as a set of elementary statements by making use of the notion of the Tschirnhaus transform $T_K(J)$ of a polynomial J by a polynomial K such that degree (K) $<$ degree (J) (see e.g. Perron [13, Section 53]):

We may replace the condition (ii) in the definition of a formally p-adic field V by

(ii)′ If $L \in V[X]$ is a monic irreducible polynomial of degree n over V then there exists $J \in M(n)$ and $K \in V[X]$ such that $\deg(K) < \deg(J)$ and $L = T_K(J)$, for all positive integers n.

Now we come to the proof of the completeness of the axioms. Let F and F' be two formally p-adic fields. We wish to show that an elementary statement holds in F if and only if it holds in F', i.e. that F and F' are *elementarily equivalent*. (We denote this relation by $F \equiv F'$.) Let F^* (resp.: F'^*) be a countable non-principal ultrapower of F (resp.: F'). Then it follows from property (**H**) preceding Theorem 5 of [1] that $F \equiv F^*$ (resp.: $F' \equiv F'^*$). In particular, this means that F^* and F'^* are formally

p-adic fields. If we assume that $|F|, |F'| \leqq 2^{\aleph_0}$, then $|F^*| = |F'^*| = 2^{\aleph_0}$. Also, F^* and F'^* are ω-pseudo-complete by Lemma 9 of [1]. Now assume the continuum hypothesis. Then we have verified that if $|F|, |F'| \leqq 2^{\aleph_0}$, then F^* and F'^* satisfy the hypothesis of Theorem 1, except for showing that $\operatorname{ord}(F^*) = \operatorname{ord}(F'^*)$. Now since $\operatorname{ord}(F)$ and $\operatorname{ord}(F')$ are Z-groups, it follows that $\operatorname{ord}(F) \equiv \operatorname{ord}(F')$. (See [15, Th. 4.4]. Z-groups are there referred to as regularly discrete groups.) Hence, by [9, Cor. A.6], $\operatorname{ord}(F^*) = \operatorname{ord}(F)^* \approx \operatorname{ord}(F')^* = \operatorname{ord}(F'^*)$. Now by identifying $\operatorname{ord}(F^*)$ with $\operatorname{ord}(F'^*)$ we may apply Theorem 1 to obtain that F^* and F'^* are analytically isomorphic, and hence that $F \equiv F'$. The cardinality restrictions on F and F' may now be dropped by a use of the Löwenheim-Skolem Theorem. Hence, we have that all formally p-adic fields are elementarily equivalent, assuming the continuum hypothesis. The assumption of the continuum hypothesis may now be eliminated as in [10, p. 235]. We have thus proved the following result.

THEOREM 2. *The elementary theory of formally p-adic fields is complete.*

Since the axioms we have given for formally p-adic fields form a recursive set, we obtain

THEOREM 3. *The elementary theory of Q_p is decidable.*

For the terminology used in the following theorem the reader is referred to [10, p. 236]. The proof is similar to the proof of Theorem 2, using in this case the model-completeness of the elementary theory of Z-groups [15, Th. 4.4].

THEOREM 4. *The elementary theory T of formally p-adic fields is model-complete.*[3]

It follows from this result via Theorem 2.4.2 [14] that every elementary formula is provably equivalent in T to an existential formula (i.e. a formula of the form $\exists x_1 \exists x_2 \cdots \exists x_n Y$, where Y is quantifier-free). The question of whether every elementary formula is equivalent to a quantifier-free formula is still unanswered.[4] This question is closely related to the problem mentioned in the introduction of whether there exists a "useful" decision method for the elementary theory of Q_p.

In the analogous real field situation the "useful" decision procedure

[3] Since, by Lemma 2, the valued field A_p is a prime model for the class of formally p-adic fields, Theorem 4 implies Theorem 2.

[4] An affirmative answer to this question is given in [2, Section 3].

is based upon Sturm's Theorem—an algorithm for determining how many real roots a polynomial with real coefficients has in an interval [17, Section 68]. An essential feature of Sturm's Theorem is that the number of steps used for determining the number of real roots is a function only of the degree of the polynomial. Does there exist a p-adic analogue of Sturm's Theorem? [4]

CORNELL UNIVERSITY.

REFERENCES.

[1] J. Ax and S. Kochen, "Diophantine problems over local fields I," *American Journal of Mathematics*, this volume.

[2] ——— and S. Kochen, "Diophantine problems over local fields III," to appear in *Annals of Mathematics*.

[3] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press, 1956.

[4] K. Gödel, *The consistency of the continuum hypothesis*, Annals of Mathematics Studies, Princeton University Press, 1940.

[5] H. Hasse, *Zahlentheorie*, Akademie-Verlag, Berlin, 2nd ed., 1963.

[6] N. Jacobson, *Lectures in Abstract Algebra III*, Van Nostrand, 1964.

[7] I. Kaplansky, "Maximal fields with valuations," *Duke Mathematical Journal*, vol. 9 (1942), pp. 313-321.

[8] ———, "Maximal fields with valuations, II," *Duke Mathematical Journal*, vol. 12 (1945), pp. 243-248.

[9] H. J. Keisler, "Ultraproducts and elementary classes," *Indagationes Mathematicae*, vol. 23 (1961), pp. 477-495.

[10] S. Kochen, "Ultraproducts in the theory of models," *Annals of Mathematics* (2), vol. 74 (1961), pp. 221-261.

[11] W. Krull, "Allgemeine Bewertungstheorie," *Journal für die Reine und Angewandte Mathematik*, vol. 167 (1932), pp. 160-196.

[12] R. Nunke, "Modules of extensions over Dedekind rings," *Illinois Journal of Mathematics*, vol. 3 (1959), pp. 222-241.

[13] O. Perron, *Algebra I*, Gruyter, 1951.

[14] A. Robinson, *Complete Theories*, Studies in Logic, North-Holland, 1956.

[15] A. Robinson and E. Zakin, "Elementary properties of ordered abelian groups," *Transactions of the American Mathematical Society*, vol. 96 (1960), pp. 222-236.

[16] K. Rychlik, "Zur Bewertungstheorie der algebraischen Körper," *Journal für die Reine und Angewandte Mathematik*, vol. 153 (1923), pp. 94-107.

[17] O. F. G. Schilling, *The Theory of Valuations*, Mathematical Surveys No. IV, American Mathematical Society.

[18] B. L. Van der Waerden, *Modern Algebra*, vol. 1, English Edition, Ungar, 1948.

FUNCTIONAL COHOMOLOGY OPERATIONS AND RELATIONS.

By JEAN-PIERRE MEYER.[1]

1. **Introduction.** It has long been clear that relations in the Steenrod algebra are the basis for the existence of secondary cohomology operations; Adams, in his fundamental paper [1], formalized this connection. The similar connection between relations in the cohomology of a space Z and functional cohomology operations, while implicit in Steenrod's original paper [20], did not become equally clear, at least to this author, until the work of Mahowald [9], and Peterson-Stein [15] on sphere-bundles. The relations we are discussing are of the form $\sum_{ij} z_{ij} \cup \theta_{ij}[z_j] = 0$,, where z_{ij}, z_j are cohomology classes of Z and θ_{ij} are (primary) cohomology operations. The main purpose of this paper is to develop, more or less systematically, the theory of such functional cohomology operations. As shown in [16], functional cohomology operations may be studied via two methods; first, the method originally used by Steenrod, secondly, the method of universal examples. While elegant, the second method is of maximum effectiveness only if one's knowledge of the cohomology of the universal example spaces is sufficiently extensive; by the very nature of the relations we consider, we must usually work outside the "stable range" where such information is adequate and thus our approach is mostly based on the Steenrod method. An important aspect of this method is the use of cochain-formulas; in order to consider coefficient groups of sufficient generality for applications, it was found necessary to study cochain operations, and their relationship to cohomology operations, in some detail; this is done in Chapter I. Next, in Chapter II, we define our functional cohomology operations, much as Steenrod did, and we prove the existence of universal examples for the operations. Although our treatment is also strongly influenced by Adams' work, [1], we are unable to characterize our operations by axioms, as he did for secondary operations. This is due to the lack of sufficient information on the cohomology of fibre spaces outside the stable range, already alluded to. In Chapter III we find that the relations we have been discussing are not

Received August 21, 1964.

[1] This work was supported in part by the U. S. Army Research Office under grant DA-31-124-ARO(D)-176.

excessively general, as they are necessary for the description of natural phenomena. More precisely, we show that every cohomology class in the total space X of a fibration, in the stable range, arises out of a unique relation in the cohomology of the base. Since this is the case, it would be useful to determine this relation from intrinsic properties of the cohomology class. We are able to do this in the special but important case of a principal fibration and find that the relation in $H^*(B)$ from which $v \in H^*(X)$ arises can be read off from $\mu^*(v)$ where $\mu\colon F \times X \to X$ is the operation of the fiber. In view of the well-known connection between such map μ and the obstruction theory of fiber-spaces, this fact will be useful in applications, as will be seen in Chapter V. In Chapter IV, various properties of functional cohomology operations are studied, which we call the naturality, additivity and product formulas. At this stage of the exposition, we may describe the latter two as stating how the functional cohomology operations behave when the relations from which they arise are composed in various ways and when the arguments are changed. Also, we make a preliminary attempt to formalize the connection between relations and operations, following [1], when all coefficient groups are isomorphic to a field. It involves the introduction of a concept, independently developed by Massey and Peterson in [10], the semi-tensor product of two algebras. Finally, in Chapter V, some applications are given; they are of two kinds: (1) using the additivity and product formulas, we obtain results on secondary characteristic classes of sphere-bundles, [18], (2) we obtain results on Moore-Postnikov systems of fibrations, the main one being that under certain stability hypotheses, each k-invariant arises out of a well-defined relation on the preceding k-invariant. Thus the indeterminacy of the obstruction defined by a certain k-invariant can be read off from a relation on the preceding k-invariant.

There is some overlap between the contents of this paper and recent work of Adem and Kristensen; these connections will be discussed at the end of Chapter V.

Chapter I. Cochain Operations.

In this chapter, we develop those parts of the theory of cochain operations which we will need in later chapters.

2. Definitions and preliminaries.

(2.1) *Definition.* Let n_i, q be positive integers, π, π_i abelian groups $(i = 1, 2, \cdots, r)$.

(a) A *cohomology operation* of type $(\{n_i\}, q; \{\pi_i\}, \pi)$ is a natural transformation of functors (from CSS-complexes to sets)

$$\theta(\cdot): \mathop{\times}_{i=1}^{r} H^{n_i}(\cdot\,; \pi_i) \to H^q(\cdot\,; \pi)$$

(b) A *cochain operation* of type $(\{n_i\}, q; \{\pi_i\}, \pi)$ is a natural transformation of functors (from CSS-complexes to sets)

$$\phi(\cdot): \mathop{\times}_{i=1}^{r} C^{n_i}(\cdot\,; \pi_i) \to C^q(\cdot\,; \pi)$$

Here C^n represents the *normalized* cochain functor. Note that we do not require θ and ϕ to be homomorphisms. Occasionally, when either the dimensional range or the coefficients are clear, we will abbreviate the type of our operation to $(\{n_i\}, q)$ or $(\{\pi_i\}, \pi)$.

There is, as is well-known, a close connection between such operations and Eilenberg-MacLane complexes, which we shall exploit to obtain properties of these operations. Let us begin by recalling certain facts and definitions, [2], [5]:

$K(\pi, n)$, resp. $L(\pi, n+1)$, is the CSS-complex whose q-cells are the elements of $Z^n(\Delta_q; \pi)$, resp. $C^n(\Delta_q; \pi)$, where Δ_q is the "standard" q-simplex, and with the obvious semi-simplicial operators, (see above references for more details).

The exact sequence

$$0 \to Z^n(\Delta_q; \pi) \to C^n(\Delta_q; \pi) \xrightarrow{\delta} Z^{n+1}(\Delta_q; \pi) \to 0$$

defines CSS-maps

$$K(\pi, n) \xrightarrow{i} L(\pi, n+1) \xrightarrow{p} K(\pi, n+1)$$

with pi the map onto the subcomplex $K(0, n+1)$ consisting of the unique zero degenerate cell in each dimension. The CSS-complex $L(\pi, n+1)$ is acyclic; indeed, a contracting homotopy $\epsilon: L_q(\pi, n+1) \to L_{q+1}(\pi, n+1)$ is defined by $\epsilon: C^n(\Delta_q; \pi) \to C^n(\Delta_{q+1}; \pi)$, where

$$\epsilon(u)(m_0, \cdots, m_n) = \begin{cases} 0 & \text{if } m_0 = 0 \\ u(m_0 - 1, \cdots, m_n - 1) & \text{if } m_0 > 0. \end{cases}$$

It is easily verified that $\epsilon d + d\epsilon = 1$. The maps

$$\sigma: L_q(\pi, n) \to K_{q+1}(\pi, n)$$
$$\Sigma: K_q(\pi, n-1) \to K_{q+1}(\pi, n)$$

are defined by $\sigma = p\epsilon$, $\Sigma = \sigma i$. They satisfy the relations

$$\sigma d + d\sigma = p \qquad \qquad (2.2)$$
$$\Sigma d + d\Sigma = 0$$

in the *normalized* complex, since the image of pi consists of degenerate simplices. By passing to cochains, $\sigma^{\#}$ and $\Sigma^{\#}$ are defined, satisfying the relations dual to (2.2); in particular, $\Sigma^{\#}$ anti-commutes with the coboundary and thus induces the *suspension* homomorphism

$$\Sigma^{*}: H^{q+1}(K(\pi, n); \pi') \to H^{q}(K(\pi, n-1); \pi').$$

The complex $L(\pi, n+1)$ has a *basic cochain* $c_n \in C^n(L(\pi, n+1); \pi)$ which assigns to each n-simplex, i.e., each element of $C^n(\Delta_n; \pi)$ the corresponding element of π, and similarly $K(\pi, n)$ has a *basis cocycle* $b_n \in Z^n(K(\pi, n); \pi)$. It is easily verified that

$$i^{\#}(c_n) = b_n, p^{\#}(b_n) = \delta c_{n-1}, \sigma^{\#}(b_n) = c_{n-1}, \Sigma^{\#}(b_n) = b_{n-1}. \qquad (2.3)$$

Let X, Y be CSS-complexes and $\mathrm{Hom}(X, Y)$ the set of CSS-maps $X \to Y$. Then the assignment $f \to f^{\#}(c_n)$, $g \to g^{\#}(b_n)$ defines natural one-one correspondences between $\mathrm{Hom}(X, L(\pi, n+1))$ and $C^n(X; \pi)$, and $\mathrm{Hom}(X, K(\pi, n))$ and $Z^n(X; \pi)$. The maps i, p induce

$$\mathrm{Hom}(X, K(\pi, n) \xrightarrow{i_{\#}} \mathrm{Hom}(X, L(\pi, n+1) \xrightarrow{p_{\#}} \mathrm{Hom}(X, K(\pi, n+1))$$

which, under the above correspondence, become

$$Z^n(X; \pi) \to C^n(X; \pi) \xrightarrow{\delta} Z^{n+1}(X; \pi).$$

If $c \in C^n(X; \pi)$, $z \in Z^n(X; \pi)$, the corresponding elements of

$$\mathrm{Hom}(X, L(\pi, n+1)), \quad \mathrm{Hom}(X, K(\pi, n))$$

will be denoted by $U(c)$, $T(z)$, respectively. Then

$$iT(Z) = U(z) \quad \text{and} \quad pU(c) = T(\delta c). \qquad (2.4)$$

Finally, the maps $T(z)$, $T(z')$ are chain-homotopic if and only if $z \backsim z'$.

We are now in a position to define the various operations which we need. Let f, g be the maps, given by the Eilenberg-Zilber theorem, [5], II, § 2, from the Cartesian product to the tensor product of complexes, and vice-versa. Let K be a CSS-complex, $z_i \in Z^{n_i}(K; \pi_i)$, $c_i \in C^{n_i}(K; \pi_i)$. Then we define chain-maps

$$T(z_1,\cdots,z_r),T'(z_1,\cdots,z_r): K\to \underset{i=1}{\overset{r}{\times}} K(\pi_i,n_i)$$

$$U(c_1,\cdots,c_r),U'(c_1,\cdots,c_r): K\to \underset{i=1}{\overset{r}{\times}} L(\pi_i,n_i+1)$$

by
$$T(z_1,\cdots,z_r)=(\underset{i=1}{\overset{r}{\times}} T(z_i))\Delta,\quad T'(z_1,\cdots,z_r)=g(\bigotimes_{i=1}^{r} T(z_i))f\Delta$$

$$U(c_1,\cdots,c_r)=(\underset{i=1}{\overset{r}{\times}} U(c_i))\Delta,\quad U'(c_1,\cdots,c_r)=g(\bigotimes_{i=1}^{r} U(c_i))f\Delta$$

where $\Delta: K\to K^r$ is the diagonal map. Clearly, if $z_i \smile z'_i$, $i=1,\cdots,r$, then the pairs of maps $T(z_1,\cdots,z_r)$, $T(z'_1,\cdots,z'_r)$ and $T'(z_1,\cdots,z_r)$, $T'(z'_1,\cdots,z'_r)$ are chain homotopic.

If $h\in H^q(\underset{i=1}{\overset{r}{\times}} K(\pi_i,n_i);\pi)$, $c\in C^q(\underset{i=1}{\overset{r}{\times}} L(\pi_i,n_i+1);\pi)$, define

$$\begin{aligned}\theta(h)[\{z_1\},\cdots,\{z_r\}]&=T(z_1,\cdots,z_r)^*(h)\\ \theta'(h)[\{z_1\},\cdots,\{z_r\}]&=T'(z_1,\cdots,z_r)^*(h)\\ \phi(c)[c_1,\cdots,c_r]&=U(c_1,\cdots,c_r)^\#(c)\\ \phi'(c)[c_1,\cdots,c_r]&=U'(c_1,\cdots,c_r)^\#(c).\end{aligned}\tag{2.5}$$

In view of the last remark of the preceding paragraph, $\theta(h)$ and $\theta'(h)$ are well-defined. It is immediate that $\theta(h)$, $\theta'(h)$ are cohomology operations of type $(\{n_i\},q;\{\pi_i\},\pi)$ and that $\phi(c)$, $\phi'(c)$ are cochain operations of the same type. Conversely, if θ, ϕ are cohomology and cochain operations of the above type, then it is not hard to see that $\theta=\theta(h)$, $\phi=\phi(c)$, where $h=\theta[p_1{}^*b_{n_1},\cdots,p_r{}^*b_{n_r}]$, $c=\phi[p_1{}^\#c_{n_1},\cdots,p_r{}^\#c_{n_r}]$, and p_i is the projection onto the i-th factor.

We now define the *suspension* of cohomology and cochain operations. Since we are dealing with operations on several variables, we "suspend" with respect to one variable at a time: let s be an integer, $1\leqq s\leqq r$, and $\theta=\theta(h)$, $\phi=\phi(c)$, be as above. The *s-suspension* of θ, ϕ, denoted by $\Sigma_s(\theta)$, $\sigma_s(\phi)$, respectively, are operations of type $(\{n_i-\delta_{si}\},q-1;\{\pi_i\},\pi)$, where δ_{si} is the Kronecker delta-symbol, given by the equations

$$\begin{aligned}\Sigma_s(\theta)[\{z_1\},\cdots,\{z_r\}]&=[g(\bigotimes_i \Sigma^{\delta_{si}})(\bigotimes_i T(z_i))f\Delta]^*(h)\\ \sigma_s(\phi)[c_1,\cdots,c_r]&=[g(\bigotimes_i i\sigma^{\delta_{si}})(\bigotimes_i U(c_i))f\Delta]^\#(c).\end{aligned}\tag{2.6}$$

If $r=1$, we write Σ, σ for Σ_1, σ_1.

3. Properties of special cochain operations. We now investigate the connection between cochain and cohomology operations.

(3.1) *Definition.* A cochain operation ϕ is *special* if $\delta\phi[c_1,\cdots,c_r] = 0$ whenever $\delta c_1 = 0,\cdots,\delta c_r = 0$.

Let $j\colon \bigtimes_{i=1}^{r} K(\pi_i, n_i) \to \bigtimes_{i=1}^{r} L(\pi_i, n_i+1)$ denote the inclusion.

(3.2) PROPOSITION. *$\phi = \phi(c)$ is special if and only if $j^{\#}\delta c = 0$; $\phi'(c)$ is special if $j^{\#}\delta c = 0$.*

Proof. Let $\delta c_1 = 0,\cdots,\delta c_r = 0$. Then

$$\begin{aligned}\delta\phi(c)[c_1,\cdots,c_r] &= \delta U(c_1,\cdots,c_r)^{\#}(c) = \delta T(c_1,\cdots,c_r)^{\#}j^{\#}(c)\\ &= T(c_1,\cdots,c_r)^{\#}j^{\#}\delta c\end{aligned}$$

and $\delta\phi'(c)[c_1,\cdots,c_r] = T'(c_1,\cdots,c_r)^{\#}j^{\#}\delta c$. Thus if $j^{\#}\delta c = 0$, then $\phi(c)$, $\phi'(c)$ are special. Conversely, let $\phi(c)$ be special, $K = \bigtimes_{i=1}^{r} K(\pi_i, n_i)$, $c_i = p_i{}^{\#}b_{n_i}$. Then $T(c_1,\cdots,c_r) =$ identity and it follows that $j^{\#}\delta c = 0$. Note that the last argument does not apply to $\phi'(c)$, and that $\phi'(c)$ is special if $\phi(c)$ is special.

(3.3) PROPOSITION. *$\phi'(c)[c_1,\cdots,c_r]$, and therefore $\theta'(h)[h_1,\cdots,h_r]$, vanishes whenever one of the arguments is zero.*

Proof. If $c_s = 0$, then $U(c_s)$ is the map onto the subcomplex $L(0, n_s+1)$ constisting of the base-vertex and its degeneracies. Therefore the image of $U'(c_1,\cdots,c_r)$ consist of degenerate chains (Theorem 2.1a, [5], II) and $\phi'(c)[c_1,\cdots,c_r] = U'(c_1,\cdots,c_r)^{\#}(c) = 0$, since c is a normalized cochain.

(3.4) PROPOSITION. $\delta\sigma_s(\phi(c))[c_1,\cdots,c_r] = \phi'(c)[c_1,\cdots,\delta c_s,\cdots,c_r] - \sigma_s(\phi(\delta c))[c_1,\cdots,c_r]$

Proof.
$$\begin{aligned}\delta\sigma_s(\phi(c))[c_1,\cdots,c_r] &= \delta\Delta^{\#}f^{\#}(\textstyle\bigotimes_i U(c_i))^{\#}(\bigotimes_i (i\sigma)^{\delta_{si}})^{\#}g^{\#}(c)\\ &= \Delta^{\#}f^{\#}(\textstyle\bigotimes_i U(c_i))^{\#}\delta(\bigotimes_i (i\sigma)^{\delta_{si}})^{\#}g^{\#}(c)\\ &= \Delta^{\#}f^{\#}(\textstyle\bigotimes_i U(c_i))^{\#}[-(\bigotimes_i (i\sigma)^{\delta_{si}})^{\#}\delta\\ &\qquad + (\textstyle\bigotimes_i (ip)^{\delta_{si}})^{\#}]g^{\#}(c) \text{ by } (2.2)\\ &= -\sigma_s(\phi(\delta c))[c_1,\cdots,c_r] + \Delta^{\#}f^{\#}(\textstyle\bigotimes_i U(d_i))^{\#}g^{\#}(c)\end{aligned}$$

where $d_i = c_i$, $i \neq s$ and $d_s = \delta c_s$ since $ipU(c_s) = iT(\delta c_s) = U(\delta c_s)$ by (2.4). The result is now immediate.

(3.5) COROLLARY. *If $r = 1$ and ϕ is special, then $\delta\sigma(\phi)[c_1] = \phi[\delta c_1]$.*

Proof. Let $\phi = \phi(c) = \phi'(c)$. Then, if ϕ is special we have $j^{\#}\delta c = 0$ by (3.2) and so $\sigma(\phi(\delta c)) = \phi(\sigma^{\#}j^{\#}\delta c) = \phi(0) = 0$. The conclusion now follows from (3.4).

(3.6) COROLLARY. *If $\phi(c)$ is special, then so is $\sigma_s(\phi(c))$.*

Proof. Let $\delta c_1 = 0, \cdots, \delta c_r = 0$. Then, in particular, $\delta c_s = 0$ for some s, $1 \leqq s \leqq r$, and $\phi'(c)[c_1, \cdots, \delta c_s, \cdots, c_r] = 0$ by (3.3). On the other hand

$$\begin{aligned}\sigma_s(\phi(\delta c))[c_1, \cdots, c_r] &= [g(\otimes(i\sigma)^{\delta_{si}})(\otimes U(c_i))f\Delta]^{\#}(\delta c)\\ &= [g(\otimes(i\sigma)^{\delta_{si}})(\otimes iT(c_i))f\Delta]^{\#}(\delta c)\\ &= [g(\otimes i)(\otimes(\sigma i)^{\delta_{si}})(\otimes T(c_i))f\Delta]^{\#}\delta c\\ &= [jg(\otimes(\sigma i)^{\delta_{si}})(\otimes T(c_i))f\Delta]^{\#}\delta c\\ &= 0 \text{ since } j^{\#}\delta c = 0 \text{ by (3.2).}\end{aligned}$$

The result now follows from (3.4).

The cochain operations $\phi(c)$ and $\phi'(c)$ are clearly closely related. The next proposition clarifies this connection.

(3.7) PROPOSITION. *Let $\Phi\colon \underset{i}{\times} L(\pi_i, n_i + 1) \to \underset{i}{\times} L(\pi_i, n_i + 1)$ be the chain-homotopy between gf and the identity map, then*

$$\begin{aligned}&\phi'(c)[c_1, \cdots, c_r] - \phi(c)[c_1, \cdots, c_r]\\ &\qquad = \phi(\Phi^{\#}\delta c)[c_1, \cdots, c_r] + \delta\phi(\Phi^{\#}c)[c_1, \cdots, c_r]\end{aligned}$$

In particular, if $\phi(c)$ is special and if $\delta c_1 = 0, \cdots, \delta c_r = 0$, then

$$\phi'(c)[c_1, \cdots, c_r] - \phi(c)[c_1, \cdots, c_r] = \delta\phi(\Phi^{\#}c)[c_1, \cdots, c_r].$$

Proof. $\phi'(c)[c_1, \cdots, c_r] - \phi(c)[c_1, \cdots, c_r]$

$$\begin{aligned}&= \Delta^{\#}f^{\#}(\otimes U(c_i))^{\#}g^{\#}(c) - \Delta^{\#}(\times U(c_i))^{\#}(c)\\ &= \Delta^{\#}(\times U(c_i))^{\#}f^{\#}g^{\#}(c) - \Delta^{\#}(\times U(c_i))^{\#}(c)\\ &= \Delta^{\#}(\times U(c_i))^{\#}[\delta\Phi^{\#} + \Phi^{\#}\delta]c\\ &= \delta\phi(\Phi^{\#}c)[c_1, \cdots, c_r] + \phi(\Phi^{\#}\delta c)[c_1, \cdots, c_r].\end{aligned}$$

If $\phi(c)$ is special, then $j^{\#}\delta c = 0$; it now follows from the naturality of Φ that if $\delta c_1 = 0, \cdots, \delta c_r = 0$, then $\phi(\Phi^{\#}\delta c)[c_1, \cdots, c_r] = 0$. Thus the last formula of (3.7) is proved.

We can now state the main theorem of this section.

(3.8) THEOREM. *If $\phi = \phi(c)$ is special, then $\phi(c)$ and $\phi'(c)$ define one and the same cohomology operation θ of the same type, by the equations*

$$\begin{aligned}\theta[\{z_1\},\cdots,\{z_r\}] &= \{\phi[z_1,\cdots,z_r]\}\\ &= \{\phi'[z_1,\cdots,z_r]\}.\end{aligned}$$

Proof. We have already seen, in the preceding proposition, that $\phi[z_1,\cdots,z_r] \smile \phi'[z_1,\cdots,z_r]$. It remains therefore to show that if $z_i \smile z'_i$, $i = 1,\cdots,r$, then $\phi[z_1,\cdots,z_r] \smile \phi[z'_1,\cdots,z'_r]$.

We have $\phi[z_1,\cdots,z_r] = T(z_1,\cdots,z_r)^{\#}j^{\#}(c)$ and $\phi[z'_1,\cdots,z'_r] = T(z'_1,\cdots,z'_r)^{\#}j^{\#}(c)$. Since $z_i \smile z'_i$, $i = 1,\cdots,r$, we see that $T(z_i)$, $T(z'_i)$ are chain-homotopic, $i = 1,\cdots,r$, and therefore that $T(z_1,\cdots,z_r)$, $T(z'_1,\cdots,z'_r)$ are chain-homotopic. Thus there is a chain-homotopy D such that

$$\begin{aligned}T(z_1,\cdots,z_r)^{\#}j^{\#}(c) - T(z'_1,\cdots,z'_r)^{\#}j^{\#}(c) &= \delta Dj^{\#}(c) + D\delta j^{\#}(c)\\ &= \delta Dj^{\#}(c)\end{aligned}$$

since $\phi(c)$ is special. It follows that $\phi[z_1,\cdots,z_r] \smile \phi[z'_1,\cdots,z'_r]$.

We will describe the conclusion of the above theorem by saying that ϕ (or ϕ') is a *representative cochain-operation* for θ.

The next theorem is a sort of converse of (3.8).

(3.9) THEOREM. *If θ is a cohomology operation of type $(\{n_i\}, q; \{\pi_i\}, \pi)$ then there exists a special cochain operation ϕ of the same type representing it. If $\bar{\phi}$ is another such cochain operation, then there exist cochain operations ζ and η, of type $(\{n_i\}, q-1; \{\pi_i\}, \pi)$ and $(\{n_i\}, q; \{\pi_i\}, \pi)$ respectively such that*

$$\phi - \bar{\phi} = \delta\zeta + \eta$$

where $\eta[c_1,\cdots,c_r] = 0$ whenever $\delta c_1 = 0,\cdots,\delta c_r = 0$.

Proof. We have already seen that $\theta = \theta(h)$ where $h = \theta[p_1^*\{b_{n_1}\},\cdots, p_r^*\{b_{n_r}\}] \in H^q(\mathop{\times}\limits_{i=1}^{r} K(\pi_i, n_i); \pi')$. Let $z \in Z^q(\mathop{\times}\limits_{i=1}^{r} K(\pi_i; n_i); \pi')$ be a representative cocycle of h and extend z to a cochain $c \in C^q(\mathop{\times}\limits_{i=1}^{r} L(\pi_i, n_i + 1); \pi')$. Then $\phi = \phi(c)$ is a representative cochain operation for θ. If $\bar{\phi} = \phi(c')$ is another representative cochain operation, then $j^{\#}(c - c') = \delta d$. Extend d to a cochain d' in $\mathop{\times}\limits_{i=1}^{r} L(\pi_i, n_i + 1)$. Then $c - c' = \delta d' + e$ where $j^{\#}(e) = 0$, and therefore $\phi - \bar{\phi} = \delta\phi(d') + \phi(e)$. Letting $\zeta = \phi(d')$, $\eta = \phi(e)$, the theorem is proved.

(3.10) *Remark.* It would be tempting (and useful!) to assume that η in (3.9) has the form $\eta'\delta$, i.e. that $\eta[c_1,\cdots,c_r]=\eta'[\delta c_1,\cdots,\delta c_r]$. This, however, is false in general, although true in the "stable case."

In a later computation we will need, in a special case, somewhat more detailed information than that contained in (3.8). The next proposition provides this information.

(3.11) PROPOSITION. *Let ϕ be a special cochain operation of type $(n,\pi;q;\pi')$, $a\in Z^n(K;\pi)$, $b\in C^{n-1}(K;\pi)$. Then there is a cochain operation ϕ_1 of type $(\{n,n-1\},q-1;\{\pi,\pi\},\pi')$ such that*

$$\phi[a-\delta b]=\phi[a]-\phi[\delta b]+\delta\phi_1[a,b].$$

Furthermore, $\phi_1[\delta b,b]\smile 0$.

Proof. Consider the "universal examples" $L=L(\pi,n+1)\times L(\pi,n)$, $K=K(\pi,n)\times L(\pi,n)$ and the inclusion $k\colon K\subset L$. Let $\psi\in C^q(L;\pi')$ be defined by $\psi=\phi(c)[\pi_1{}^{\#}c_n-\delta\pi_2{}^{\#}c_{n-1}]-\phi(c)[\pi_1{}^{\#}c_n]+\phi(c)[\delta\pi_2{}^{\#}c_{n-1}]$. Then $k^{\#}\psi\in Z^q(K;\pi')$ since $\phi(c)$ is special, and by (3.8), $\{k^{\#}\psi\}=0$. Thus $k^{\#}\psi=\delta\bar{\zeta}$, $\bar{\zeta}\in C^{q-1}(K;\pi')$. Extend $\bar{\zeta}$ to a cochain $\zeta\in C^{q-1}(L;\pi')$; $k^{\#}(\psi-\delta\zeta)=0$ and $\psi-\delta\zeta=\eta$ where $\eta\in C^q(L;\pi')$ such that $k^{\#}(\eta)=0$. Hence $\Delta^{\#}(U(a)\times U(b))^{\#}(\psi-\delta\zeta-\eta)=0$, i.e.

$$\phi[a-\delta b]-\phi[a]+\phi[\delta b]=\delta\phi(\zeta)[a,b]+\phi(\eta)[a,b]$$

where $\phi(\eta)[a,b]=0$ if $\delta a=0$. Letting $\phi_1=\phi(\zeta)$, the first part of the proposition is proved. To prove $\phi_1[\delta b,b]\smile 0$, let $a=\delta b$ in the above formula, then $\phi[0]=\phi[\delta b]-\phi[\delta b]+\delta\phi_1[\delta b,b]$, so $\delta\phi_1[\delta b,b]=0$.

$$\begin{aligned}\text{But } \phi_1[\delta b,b]&=\Delta^{\#}(U(\delta b)\times U(b))^{\#}\zeta=\Delta^{\#}(T(\delta b)\times U(b))^{\#}\bar{\zeta}\\&=\Delta^{\#}(pU(b)\times U(b))^{\#}\bar{\zeta}=\Delta^{\#}(U(b)\times U(b))^{\#}(p\times 1)^{\#}\bar{\zeta}\\&=U(b)^{\#}\Delta^{\#}(p\times 1)^{\#}\bar{\zeta}\end{aligned}$$

so $\Delta^{\#}(p\times 1)^{\#}\bar{\zeta}\in Z^{q-1}(L(\pi,n);\pi')$. The result now follows since $L(\pi,n)$ is acyclic.

Another proposition of the same type is the following:

(3.12) PROPOSITION. *Let ϕ be a special cochain operation of type $(n+1,q+1;\pi,\pi')$, $a\in Z^n(K;\pi)$, $b\in C^n(K;\pi)$. Then there is a cochain operation η of type $(\{n,n\},q-1;\{\pi,\pi\},\pi')$ such that*

$$\sigma(\phi)[a+b]=\sigma(\phi)[a]+\sigma(\phi)[b]+\delta\eta[a,b]$$

Proof. By [5], III, Theorem 16.2, the cohomology operation $\Sigma(\theta)$ represented by $\sigma(\phi)$ is additive. Thus the 2 cochain operations $\sigma(\phi)[c_1+c_2]$

and $\sigma(\phi)[c_1]+\sigma(\phi)[c_2]$ on two variables represent the same cohomology operation (on two variables). Hence, by (3.9), there are cochain operations ζ_1, ζ_2 such that

$$\sigma(\phi)[a+b]=\sigma(\phi)[a]+\sigma(\phi)[b]+\delta\zeta_1[a,b]+\zeta_2[a,b]$$

where $\zeta_2[a,b]=0$ whenever $\delta a=0$, $\delta b=0$. Taking the coboundary of both sides, and using (3.5), we obtain

$$\phi[\delta a+\delta b]=\phi[\delta a]+\phi[\delta b]+\delta\zeta_2[a,b].$$

Let now $\delta a=0$. Then it follows that $\delta\zeta_2[a,b]=0$, whenever $\delta a=0$ and for all b. As in the proof of (3.11), the universal example for ζ_2 yields a cocycle of $(K(\pi,n)\times L(\pi,n+1),K(\pi,n)\times L(0,n+1))$. Since this pair is acyclic, ζ_2 is a coboundary and the result follows.

(3.13) COROLLARY. *If* $c\smile\sum_i c_i$ *where all but one of the* c_i*'s are cocycles, then*

$$\sigma(\phi)[c]\smile\sum_i\sigma(\phi)[c_i].$$

Proof. This follows immediately from (3.12) by induction and using (3.5).

Chapter II. Functional Cohomology Operations.

Following a general discussion of functional cohomology operations, we define in this chapter the operations in which we are primarily interested, namely those arising out of certain relations in the cohomology of a space. Our approach is as follows: (a) the operations are defined in a manner generalizing the original definition of Steenrod [20], (b) cochain formulas are obtained for representatives of the operations, (c) using the cochain formulas, the existence of universal examples for the operations is proved.

4. Definitions.

(4.1) *Definition.* S is a *natural set of mappings* (into the space Z) if it associates with each space X a subset $S(X)$ of Z^X subject to the conditions:

(a) If $f\colon X\to Y$ and $g\in S(Y)$, then $gf\in S(X)$.

(b) If $f\colon X\to Y$ is a weak homotopy-equivalence, and $gf\in S(X)$, then $g\in S(Y)$.

(4.2) *Definition.* Φ is a *functional cohomology operation* (defined on

S and with values of degree m) if Φ associates with each space X and map $h \in S(X)$ a non-empty subset $\Phi_h \subset H^m(X;G)$ (G is a fixed abelian group) subject to the conditions:

(a) If $f\colon X \to Y$ and $g \in S(Y)$, then $f^*\Phi_g \subset \Phi_{gf}$.

(b) If $f\colon X \to Y$ is a weak homotopy-equivalence and $g \in S(Y)$, then $f^*\Phi_g = \Phi_{gf}$.

(4.3) *Definition.* Let S be a natural set of mappings (into Z). Let U be a space and $p\colon U \to Z$. Then (U,p) is a ***universal example for*** S if:

(a) $p \in S(U)$.

(b) For each CW-complex X and $f \in S(X)$, there is a map $g\colon X \to U$ such that $f = pg$.
Clearly, a pair (U,p) determines a unique natural set of mappings (into Z) admitting it as universal example.

(4.4) *Definition.* Let S be a natural set of mappings, with universal example (U,p). Let Φ be a functional cohomology operation defined on S, with values of degree m; let $\mathcal{U} \subset H^m(U;G)$. Then $(U,p,\mathcal{U})$ is a ***universal example*** for Φ if:

(a) $\mathcal{U} \subset \Phi_p$.

(b) For each CW-complex X, each $f \in S(X)$ and each $y \in \Phi_f$, there is a map $g\colon X \to U$ such that $pg = f$ and $y \in g^*(\mathcal{U})$.

(c) $\mathcal{U}$ is minimal with respect to (a) and (b).
A triple $(U,p,\mathcal{U})$ determines a unique functional cohomology operation for which it serves as a universal example.

(4.5) *Definition.* A functional cohomology operation Φ is *0-ary* (*or of 0-th kind*) if:

(a) Φ_f is defined for all $f\colon X \to Z$.

(b) Φ_f is a single element of $H^m(X;G)$.
We then have the trivial proposition.

(4.6) PROPOSITION. *If Φ is a 0-ary functional cohomology operation, then it admits as universal example $(Z,1,\Phi_1)$, where 1 denotes the identity map of Z. Conversely, any element $u \in H^m(Z;G)$ defines a 0-ary functional cohomology operation, namely the one with universal example $(Z,1,u)$.*

Let Λ be a finite set of indices and $\{\phi_\lambda\}$ a Λ-tuple of 0-ary functional

cohomology operations, with universal examples $(Z,1,u_\lambda)$, $u_\lambda \in H^{m_\lambda}(Z;G_\lambda)$. We define a natural set of mappings (into Z) as follows: $T(X)$ is the set of maps $f\colon X \to Z$ such that $(\phi_\lambda)_f = f^*(u_\lambda) = 0$, for all $\lambda \in \Lambda$. Clearly T is a natural set of mappings. If Φ is a functional cohomology operation defined on such a set T, we call it a *primary functional cohomology operation.*

Let $p\colon U \to Z$ be a canonical fibration associated with $\{u_\lambda\}$, i.e., the fiber of the fibration p is a generalized Eilenberg-MacLane space and the $\{u_\lambda\}$ are the k-invariants. The following theorem is well-known (see, for example, [1]).

(4.7) THEOREM. *(U,p) is a universal example for the set T.*

5. Operations arising from relations. In the section, we wish to define a class of primary functional cohomology operations arising from relations in $H^*(Z;G)$ of the form:

$$\sum_{ij} (-1)^{m+1-n_{ij}} z_{ij} \cup \theta_{ij}[z_j] = 0 \tag{5.1}$$

where G, G_{ij}, G'_{ij} are abelian groups, $z_j \in H^{n_j}(Z;G_j)$, θ_{ij} are cohomology operations (primary) of type $(n_j, n_{ij}; G_j, G_{ij})$, $z_{ij} \in H^{m+1-n_{ij}}(X;G'_{ij})$ and the cup-product of z_{ij} and $\theta_{ij}[z_j]$ is taken with respect to a fixed pairing $G'_{ij} \otimes G_{ij} \to G$.

Let T be the natural set of mappings determined by $\{z_j\}$, and $f \in T(X)$, i.e., $f\colon X \to Z$ and $f^*(z_j) = 0$. Following Steenrod's procedure, [20], we define a subset Φ_f of $H^m(X;G)$ by considering the diagram below:

$$\begin{array}{ccccccc}
 & & \Sigma_j H^{n_j-1}(X;G_j) & \xrightarrow{\delta} & \Sigma_j H^{n_j}(Z_f,X;G_j) & \xrightarrow{i^*} & \Sigma_j H^{n_j}(Z_f;G_j) \xrightarrow{f'^*} \Sigma_j H^{n_j}(X; \\
 & & \downarrow{\scriptstyle \xi_1} & & \downarrow{\scriptstyle \xi_2} & & \downarrow{\scriptstyle \xi_3} \\
H^m(Z_f;G) & \xrightarrow{f'^*} & H^m(X;G) & \xrightarrow{\delta} & H^{m+1}(Z_f,X;G) & \xrightarrow{i^*} & H^{m+1}(Z_f;G)
\end{array}$$

Z_f denotes the mapping cylinder of f, $i\colon Z_f \to (Z_f,X)$ the inclusion, and $f'\colon X \to Z_f$ the inclusion. Recall that Z, Z_f have the same homotopy type, that (l,r) is an equivalence between them, where $l\colon Z \subset Z_f$ and $r\colon Z_f \to Z$ is the retraction, and that $rf' = f$. The vertical mappings, *not* homomorphisms in general, are given by

$$\xi_1(\{a_j\}) = \sum_{ij} f'^* z'_{ij} \cup \Sigma(\theta_{ij})[a_j],$$

$$\xi_k(\{a_j\}) = \sum_{ij} (-1)^{m+1-n_{ij}} z'_{ij} \cup \theta_{ij}[a_j], \qquad k = 2, 3,$$

where $z'_{ij} = r^*(z_{ij})$. It is easily verified that the squares are commutative. Since $f^*(z_j) = 0$, we see that $f'^*(z'_j) = 0$, where $z'_j = r^*(z_j)$. Thus, in standard fashion, there exist $z''_j \in H^{n_j}(Z_f, X; G_j)$ such that $i^*(z''_j) = z'_j$. Then $i^*\xi_2(\{z''_j\}) = \xi_3 i^*(\{z''_j\}) = \xi_3(\{z'_j\}) = \sum_{ij} (-1)^{m+1-n_{ij}} z'_{ij} \cup \theta_{ij}[z'_j] = 0$, and there exists $w \in H^m(X; G)$ such that $\delta w = \xi_2(\{z''_j\})$. The set Φ_f is defined to be the set of all w obtainable in the above fashion. To find the indeterminacy of w, let $\bar{z}''_j$ be such that $i^*(\bar{z}''_j) = z'_j$. Then

$$\bar{z}''_j = z''_j + \delta a, a_j \in H^{n_j-1}(X; G_j),$$

and, by the lemma of [19],

$$\begin{aligned} \xi_2(\{\bar{z}''_j\}) &= \sum_{ij} (-1)^{m+1-n_{ij}} z'_{ij} \cup \theta_{ij}[z''_j + \delta a_j] \\ &= \sum_{ij} (-1)^{m+1-n_{ij}} z'_{ij} \cup \theta_{ij}[z''_j] + \sum_{ij} (-1)^{m+1-n_{ij}} z'_{ij} \cup \delta\Sigma(\theta_{ij})[a_j] \\ &= \xi_2(\{z''_j\}) + \delta \sum_{ij} f'^*(z'_{ij}) \cup \Sigma(\theta_{ij})[a_j] = \xi_2(\{z''_j\} + \delta\xi_1(\{a_j\}). \end{aligned}$$

Therefore, if $\delta\bar{w} = \xi_2(\{\bar{z}''_j\})$, we see that $\bar{w} - w \in \text{image}\, \xi_1 + \text{image}\, f'^*$.

(5.2) THEOREM. *Φ is a functional cohomology operation defined on the natural set of mappings T; $\Phi_f \subset H^m(X; G)$ is a coset of the subgroup*

$$f^*H^m(Z; G) + \sum_j [\sum_i f^*(z_{ij}) \cup \Sigma(\theta_{ij})] H^{n_j-1}(X; G_j)$$

Proof. The only thing left to prove is condition (a) of (4.2): $f^*\Phi_g \subset \Phi_{gf}$. This is proved as in [20].

Note that the theorem holds without any assumptions regarding the additivity of the cohomology operations θ_{ij}. We will say that Φ is the functional cohomology operation associated with the relation (5.1).

Our next task is to find cochain-formulas for the operation Φ. We merely follow the classical method, the only difficulty being that the θ_{ij} and therefore the cochain operations representing them are not given explicitly; the results of Chapter I, however, give precisely the information required for our computations.

Let then $a_{ij} \in Z^{m+1-n_{ij}}(Z; G'_{ij})$, $a_j \in Z^{n_j}(Z; G_j)$ be representative cocycles of z_{ij}, z_j, respectively and let ϕ_{ij} be cochain operations representing θ_{ij}. Since $f^*(z_j) = 0$ and (5.1) holds, there exist $b_j \in C^{n_j-1}(X; G_j)$ and $c \in C^m(Z, G)$ such that

$$f^{\#}(a_j) = \delta b_j$$

$$\sum_{ij} (-1)^{m+1-n_{ij}} a_{ij} \cup \phi_{ij}[a_j] = \delta c.$$

Let $v = f^{\#}(c) - \sum_{ij} f^{\#}(a_{ij}) \cup \sigma(\phi_{ij})[b_j]$.

(5.3) THEOREM. *The element* $v \in C^m(X; G)$ *is a cocycle; let* $\Phi(f)$ *be the set of all* $\{v\}$ *obtainable in this fashion. Then* $\Phi(f) = \Phi_f$.

Proof. It is easy to verify that $v \in Z^m(X; G)$. We will now prove $\{v\} \in \Phi_f$, thus showing that $\Phi(f) \subset \Phi_f$. In order to do this, it suffices to exhibit the following:

(1) $a'_{ij} \in Z^{m+1-n_{ij}}(Z_f; G'_{ij})$ representing z'_{ij}

(2) $a'_j \in Z^{n_j}(Z_f; G_j)$ representing z'_j

(3) $a''_j \in Z^{n_j}(Z_f, X; G_j)$ such that $i^\#(a''_j) \backsim a'_j$

(4) $\bar{v} \in C^m(Z_f; G)$ such that $f'^\#(\bar{v}) \backsim v$ and

$$\delta\bar{v} = \sum_{ij} (-1)^{m+1-n_{ij}} a'_{ij} \cup \phi_{ij}[a''_j].$$

Let $r: Z_f \to Z$ be the retraction, $a'_{ij} = r^\# a_{ij}$, $a'_j = r^\# a_j$. Since $rf' = f$, we have $f'^\#(a'_j) = f^\#(a_j) = \delta b_j$. Extend b_j to a cochain $\bar{b}_j \in C^{n_j-1}(Z_f; G_j)$. Then $a''_j = a'_j - \delta\bar{b}_j$ and

$$\bar{v} = r^\#(c) - \sum_{ij} a'_{ij} \cup \sigma(\phi_{ij})[\bar{b}_j] + \sum_{ij} a'_{ij} \cup \phi_{ij,1}[a'_j, \bar{b}_j]$$

are the required cochains; verification of this fact, using (3.11), will be left to the reader.

We will now show that $\Phi(f) = \Phi_f$, thus completing the proof of the theorem, by showing that $\Phi(f)$ and Φ_f are cosets of the same subgroup. We must show that if $\{v\} \in \Phi(f)$ and $\alpha = f^*(\beta) + \sum_{ij} f^*(z_{ij}) \cup \Sigma(\theta_{ij})[\beta_j]$ is an arbitrary element of $f^*H^m(Z) + \sum_{ij} f^*(z_{ij}) \cup \Sigma(\theta_{ij}) H^{n_j-1}(X)$, then $\{v\} + \alpha \in \Phi(f)$. Let $\Delta \in Z^m(Z)$, $\Delta_j \in Z^{n_j-1}(X)$ be representative cocycles of β and β_j, respectively. Then $f^\#(a_j) = \delta(b_j + \Delta_j)$ and

$$\sum_{ij} (-1)^{m+1-n_{ij}} a_{ij} \cup \phi_{ij}[a_j] = \delta(c + \Delta).$$

Letting $v' = f^\#(c + \Delta) - \sum_{ij} f^\#(a_{ij}) \cup \sigma(\phi_{ij})[b_j + \Delta_j]$, it is easily seen, using (3.12), that

$$\bar{v}' - v \backsim f^\#(\Delta) + \sum_{ij} f^\#(a_{ij}) \cup \sigma(\phi_{ij})[\Delta_j],$$

which concludes the proof.

6. Universal examples. By Theorem (4.7) we know that a universal example for the natural set of mappings on which Φ is defined is the fibration $p: U \to Z$ with fiber $F = \times_j K(G_j, n_j - 1)$ and k-invariants $\{z_j\}$. In this section, we will show the existence of a universal example for the operation Φ defined in § 5.

For this purpose, we recall certain properties of $p\colon U \to Z$.

(6.1) A fibration $F \xrightarrow{i} E \xrightarrow{p} B$ is *principal* if it is equivalent to one induced from a path-space fibration, i.e., if $F = \Omega X$, $f\colon B \to X$ and p is induced from the fibration $\Omega X \to EX \to X$ by f. Such a fibration admits an operation μ of the fiber F on the total space E, which is fiber-preserving:

$$\begin{array}{ccc} F \times E & \xrightarrow{\mu} & E \\ \Big\downarrow{\scriptstyle \pi_2} & & \Big\downarrow{\scriptstyle p} \\ E & \xrightarrow{p} & B \end{array}$$

(6.2) $p\colon U \to Z$ is principal.

(6.3) $g\colon K \to Z$ admits a lifting $\tilde{g}\colon K \to U$ if and only if $g^*(z_j) = 0$, all j.

(6.4) If $\tilde{g}\colon K \to U$ is a lifting of g, $\{a_j\}$ a J-tuple with $a_j \in H^{n_j-1}(K; G_j)$ and $d\colon K \to F$ a map corresponding to $\{a_j\}$, then the composition $\tilde{g}'$, $K \xrightarrow{\Delta} K \times K \xrightarrow{d \times \tilde{g}} F \times U \xrightarrow{\mu} U$ is also a lifting of g, and every lifting of g is homotopic to one of this form.

We refer to [1], [19] for proofs of these elementary facts.

(6.5) LEMMA. *Let $F \xrightarrow{i} E \xrightarrow{p} B$ be an arbitrary principal fibration, $a \in H^*(F)$ a universally transgressive element. Then if a' is a transgression cochain for a, we have*

$$\mu^\#(a') \backsim \pi_1{}^\# i^\#(a') + \pi_2{}^\#(a')$$

where, as usual, π_i is the projection of $F \times E$ onto its i-th factor.

Proof. If the lemma is true for the universal principal F-fibration, then it is also true, by naturality, for any other. We may therefore assume that E is acyclic. Let $*$ denote any appropriate base-point. Since $\mu \mid F \times * = i$ and $\mu \mid * \times E = 1$, we may write

$$\mu^\#(a') = \pi_1{}^\# i^\#(a') + \pi_2{}^\#(a') + c_1$$

where $c_1 \in C^*(F \times E, F \vee E)$. Let $\tau(a)$, the transgression of a, be represented by the cocycle c_2 with $p^\#(c_2) = \delta a'$. Then

$$\begin{aligned} \delta\mu^\#(a') &= \pi_1{}^\# i^\#(\delta a') + \pi_2{}^\#(\delta a') + \delta c_1 \\ &= \pi_1{}^\# i^\# p^\#(c_2) + \pi_2{}^\# p^\#(c_2) + \delta c_1 \\ &= \pi_2{}^\# p^\#(c_2) + \delta c_1. \end{aligned}$$

On the other hand, $\delta\mu^\#(a') = \mu^\# p^\#(c_2) - \pi_2{}^\# p^\#(c_2)$ and therefore $\delta c_1 = 0$. Hence $\{c_1\} \in H^*(F \times E, F \vee E)$ which vanishes since E is acyclic, and so c_1 is a coboundary. This proves the lemma.

(6.6) LEMMA. *Let $u \in H^m(U; G)$ be an element of Φ_p. Then*

$$\mu^*(u) = \pi_2{}^*(u) + \sum_{ij} \pi_2{}^* p^*(z_{ij}) \cup \pi_1{}^* \Sigma(\theta_{ij})[f_j]$$

where f_j is the basic class of the j-th factor in the fiber.

Proof. By (5.3), u has a representative cocycle of the form

$$v = p^\#(c) - \sum_{ij} p^\#(a_{ij}) \cup \sigma(\phi_{ij})[b_j]$$

where $p^\#(a_j) = \delta b_j$, $\sum_{ij} (-1)^{m+1-n_{ij}} a_{ij} \cup \phi_{ij}[a_j] = \delta c$. Since z_j is, up to sign, the transgression of the j-th basic class in F, we see that b_j is a transgression cochain of a universally transgressive element and we may apply (6.5) to it. Thus

$$\mu^\#(b_j) = \pi_1{}^\# i^\#(b_j) + \pi_2{}^\#(b_j) + \delta d_j$$

and

$$\begin{aligned}\mu^\#(v) &= \mu^\# p^\#(c) - \sum_{ij} \mu^\# p^\#(a_{ij}) \cup \sigma(\phi_{ij})[\mu^\# b_j] \\ &= \pi_2{}^\# p^\#(c) - \sum_{ij} \pi_2{}^\# p^\#(a_{ij}) \cup \sigma(\phi_{ij})\,[\pi_1{}^\# i^\# b_j + \pi_2{}^\# b_j + \delta d_j].\end{aligned}$$

Now, by (3.13), since $\delta\pi_1{}^\# i^\# b_j = \pi_1{}^\# i^\# \delta b_j = \pi_1{}^\# i^\# p^\# a_j = 0$, and $\delta\delta d_j = 0$, we have

$$\sigma(\phi_{ij})[\pi_1{}^\# i^\# b_j + \pi_2{}^\# b_j + \delta d_j] \backsim \sigma(\phi_{ij})[\pi_1{}^\# i^\# b_j] + \sigma(\phi_{ij})[\pi_2{}^\# b_j]$$

and

$$\begin{aligned}\mu^\#(v) &\backsim \pi_2{}^\# p^\#(c) - \sum_{ij} \pi_2{}^\# p^\#(a_{ij}) \cup \{\pi_1{}^\# \sigma(\phi_{ij})[i^\# b_j] + \pi_2{}^\# \sigma(\phi_{ij})[b_j]\} \\ &\backsim \pi_2{}^\#(v) - \sum_{ij} \pi_2{}^\# p^\#(a_{ij}) \cup \pi_1{}^\# \sigma(\phi_{ij})[i^\# b_j]\end{aligned}$$

and the lemma is proved since $i^\#(b_j)$ represents the negative of the basic class f_j.

(6.7) THEOREM. *Let $u \in H^m(U; G)$ be an element of Φ_p. Then, if $\mathcal{U} = u + p^* H^m(Z; G)$, the triple $(U, p, \mathcal{U})$ is a universal example for Φ.*

Proof. Let $f \in T(X)$ so that $f\colon X \to Z$ satisfies $f^*(z_j) = 0$. Then, by (6.3), there is a lifting $\bar{f}$ of f. We must show that

$$\Phi_f = \{\bar{f}^*(v) \mid v \in \mathcal{U};\ \bar{f} = \text{lifting of } f\}.$$

Since Φ is a functional cohomology operation, $\bar{f}^*\Phi_p \subset \Phi_f$; so, since $\mathcal{U} \subset \Phi_p$, we have $\bar{f}^*\mathcal{U} \subset \bar{f}^*\Phi_p \subset \Phi_f$ for any lifting $\bar{f}$, and

$$\{\bar{f}^*(v) \mid v \in \mathcal{U}; \bar{f} = \text{lifting of } f\} \subset \Phi_f.$$

We will show that these two sets are actually equal by proving that they are cosets of the same subgroup, $f^*H^m(Z;G) + \sum_{ij} f^*(z_{ij}) \cup \Sigma(\theta_{ij}) H^{n_j-1}(X;G_j)$. In the first place, if $\beta \in H^m(Z;G)$, then $\bar{f}^*(u + p^*\beta) = \bar{f}^*u + \bar{f}^*p^*\beta = \bar{f}^*u + f^*\beta$. Thus it suffices to show that if $\{\beta_j\}$ is a J-tuple with $\beta_j \in H^{n_j-1}(X;G_j)$, then there exist liftings of f, $\bar{f}$ and $\bar{f}'$ such that

$$\bar{f}'^*(u) - \bar{f}^*(u) = \sum_{ij} f^*(z_{ij}) \cup \Sigma(\theta_{ij})[\beta_j]. \tag{6.8}$$

Let $\bar{f}$ be a fixed lifting of f and $\bar{f}'$ the composition $\mu(d \times \bar{f})\Delta$ where $d: X \to F$ corresponds to $\{\beta_j\}$. Then

$$\begin{aligned}
\bar{f}'^*(u) &= \Delta^*(d \times \bar{f})^*\mu^*(u) \\
&= \Delta^*(d \times \bar{f})^*\{\pi_2^*(u) + \sum_{ij} \pi_2^* p^*(z_{ij}) \cup \pi_1^*\Sigma(\theta_{ij})[f_j]\} \\
&= \bar{f}^*(u) + \sum_{ij} f^*(z_{ij}) \cup \Sigma(\theta_{ij})[\beta_j]
\end{aligned}$$

and the theorem is proved.

Chapter III. The Cohomology of Fibre Spaces and Functional Cohomology Operations.

Let $F \xrightarrow{i} E \xrightarrow{p} B$ be a fibration. In this chapter, we will prove that every cohomology class v of E in the "stable range" belongs to Φ_p where Φ is a primary functional cohomology operation associated with a certain unique relation in $H^*(B)$. Furthermore, if the fibration is principal so that an operation of F on E exists, $\mu: F \times E \to E$, then the relation may be read off from $\mu^*(v)$. This useful fact will be shown later to have important consequences. Let Λ denote an arbitrary commutative ring with unit. We will use the theory of G. Hirsch giving as references, once and for all, [4], [8].

7. An application of G. Hirsch's theory.

(7.1) Theorem. *If $\sum_{0<j<n} H^j(F;\Lambda)$, considered as a Λ-module, has a Λ-basis consisting of transgressive elements γ_k with $\tau(\gamma_k) = \beta_k$, and $v \in H^j(E;\Lambda)$, $j < n$, then there is a unique relation in $H^*(B;\Lambda)$, $\sum_k \alpha_k \cup \beta_k$*

$=0$ *such that* $v \in \Phi_p$ *where* Φ *is the primary functional cohomology operation associated with the above relation.*

Proof. Recall that, according to Hirsch, there is a monomorphism

$$u\colon C^*(B;\Lambda) \otimes H^*(F;\Lambda) \to C^*(E;\Lambda)$$

with the following properties:

(1) the image of u is a subcomplex whose cohomology is isomorphic to $H^*(E;\Lambda)$.

(2) $u(b \otimes \gamma) = p^\#(b) \cup u(1 \otimes \gamma)$.

(3) if γ is transgressive with c a transgression cochain for γ, then we may choose $u(1 \otimes \gamma) = c$.

Consider now $v \in H^j(E;\Lambda)$, $j < n$; in view of the above, v may be represented by

$$u(a \otimes 1 + \sum_k a_k \otimes \gamma_k) = p^\#(a) + \sum_k p^\#(a_k) \cup c_k$$

where c_k is a transgression cochain for γ_k, degree $(a_k) = i_k$ and $p^\#(b_k) = \delta c_k$, b_k representing $\beta_k = \tau(\gamma_k)$. Let us now express the fact that the above cochain is a cocycle:

$$\begin{aligned}
0 &= \delta(p^\#(a) + \sum_k p^\#(a_k) \cup c_k) \\
&= p^\#(\delta a) + \sum_k p^\# \delta a_k \cup c_k + \sum_k (-1)^{i_k} p^\# a_k \cup \delta c_k \\
&= p^\# \delta a + \sum_k p^\# \delta a_k \cup c_k + \sum_k (-1)^{i_k} p^\# a_k \cup p^\# b_k \\
&= p^\#[\delta a + \sum_k (-1)^{i_k} a_k \cup b_k] + \sum_k p^\# \delta a_k \cup c_k \\
&= u\{[\delta a + \sum_k (-1)^{i_k} a_k \cup b_k] \otimes 1 + \sum_k \delta a_k \otimes \gamma_k\}.
\end{aligned}$$

Since u is a monomorphism, we obtain

$$\delta a + \sum_k (-1)^{i_k} a_k \cup b_k = 0$$

$$\delta a_k = 0.$$

Thus, $v \in \Phi_p$ where Φ is the primary functional cohomology operation associated with the relation $\sum_k (-1)^{i_k - 1} \alpha_k \cup \beta_k = 0$ where $\alpha_k = \{a_k\}$. It remains to verify that this relation is unique (for a fixed choice of basis $\{\gamma_k\}$).

Let us suppose, then, that v is also represented by $u(a' \otimes 1 + \sum_k a'_k \otimes \gamma_k)$

so that degree $(a'_k) = i_k$, $\delta a'_k = 0$ and $\delta a' + \sum_k (-1)^{i_k} a'_k \cup b_k = 0$. Then there exist d, d_k such that

$$\delta u(d \otimes 1 + \sum_k d_k \otimes \gamma_k) = u(a \otimes 1 + \sum_k a_k \otimes \gamma_k) - u(a' \otimes 1 + \sum_k a'_k \otimes \gamma_k).$$

It follows easily from this that $\delta d_k = a_k - a'_k$ so that $\{a_k\} = \{a'_k\}$ and the relation $\sum_k (-1)^{i_k - 1} \alpha_k \cup \beta_k = 0$ is uniquely determined.

(7.3) PROPOSITION. *Under the same hypothesis as* (7.1), *if* $j < n$, *then kernel* $p^* \cap H^j(B; \Lambda)$ *is the subset of elements of degree* j *in the ideal generated by* $\{\beta_k\}$.

Proof. Let $\beta = \{b\} \in$ kernel $p^* \cap H^j(B; \Lambda)$, $j < n$. Then there exist a, a_k such that

$$p^\#(b) = u(b \otimes 1) = \delta u(a \otimes 1 + \sum_k a_k \otimes \gamma_k)$$
$$= u\{[\delta a + \sum_k \pm a_k \cup b_k] \otimes 1 + \sum_k \delta a_k \otimes \gamma_k\}.$$

Hence $b = \delta a + \sum \pm a_k \cup b_k$, $\delta a_k = 0$ and the result follows.

(7.4) THEOREM. *Let* $F \xrightarrow{i} E \xrightarrow{p} B$ *be a principal fibration with operation* $\mu\colon F \times E \to E$ *and assume that the basis* $\{\gamma_k\}$ *of* $\sum_{0<j<n} H^j(F; \Lambda)$ *consists of universally transgressive elements. Then the relation* $\sum_k \alpha_k \cup \beta_k = 0$ *associated with* $v \in H^j(E; \Lambda)$, $j < n$ *and* $\mu^*(v)$ *are connected by the equation*

$$\mu^*(v) = \pi_2^*(v) + \sum (-1)^{i_k} \pi_2^* p^*(\alpha_k) \cup \pi_1^*(\gamma_k)$$

Proof. The proof is essentially the same as that of Lemma (6.6).

(7.5) *Example.* Let $F = S^k$. Then we may take $n = \infty$; $\{\gamma_k\}$ consists of the single element $S =$ generator of $H^k(S^k; \Lambda)$ and every element of $H^*(E; \Lambda)$ is associated with a relation of the form $\beta \cup W_{k+1} = 0$, where $W_{k+1} = \tau(S)$ is the characteristic class of the fibration. The kernel of p^* is the ideal generated by W_{k+1}. These facts, taken together, are of course equivalent to the information contained in the Gysin sequence of p.

(7.6) *Example.* Let $F = \times_i K(\Lambda, n_i)$. Then we may take $n = 2\nu$ where $\nu = \min_i n_i$; $\{\gamma_k\}$ consists of elements $\theta_{ij}[b_i]$ where b_i is the basic class (of degree n_i) corresponding to the i-th factor $K(\Lambda, n_i)$ and θ_{ij} are cohomology

operation of type $(n_i, n_{ij}, \Lambda, \Lambda)$, $n_{ij} < 2\nu \leqq 2n_i$. Thus $\theta_{ij} = \Sigma(\psi_{ij})$ where ψ_{ij} is of type $(n_i + 1, n_{ij} + 1, \Lambda, \Lambda)$ and every element $v \in H^j(E;\Lambda)$, $j < 2\nu$ is associated with a unique relation of the type $\sum_{ij} z_{ij} \cup \tau(\theta_{ij}[b_i]) = 0$ i.e.,

$$\sum_{ij} z_{ij} \cup \psi_{ij}[z_i] = 0 \tag{7.7}$$

where $z_i = \tau(b_i)$ is the i-th k-invariant of the fibration. Note, furthermore, that degree $z_{ij} = j + 1 - (n_{ij} + 1) < \nu$ and in such dimensions, p^* is a monomorphism. Thus, by (7.4), the relation (7.7) can be read off from $\mu^*(v)$.

Chapter IV. Properties of Functional Cohomology Operations.

In this chapter, we prove three important properties of the functional cohomology operations introduced earlier; we will call these the "naturality," "linearity" and "product" formulas. It should be noted that condition (b) of (4.2), the definition of a functional cohomology operation, is already a naturality property, namely "naturality with respect to the domain space." We now wish to consider "naturality at the target," so to speak. The theorems of this chapter could be proved by either of two methods: (i) the diagram definition of the operations, (ii) the cochain definition of the operations. Rather than follow a consistent approach, we have used in each case what seemed the easiest method.

It often happens in the study of cohomology operations that one wishes to compare two operations whose indeterminates, $\mathrm{ind}(\Phi)$ and $\mathrm{ind}(\Psi)$, while not strictly comparable, nevertheless intersect in a non-empty common subset. The conclusion often is that $\Phi \cap \Psi \neq \phi$. An equivalent statement is

$$\Phi = \Psi \text{ modulo } \mathrm{ind}(\Phi) + \mathrm{ind}(\Psi),$$

the so-called "total indeterminacy" of the formula, or of the two operations. We shall use these two statements interchangeably.

8. Naturality.

(8.1). THEOREM. *Let* $\sum_{ij} (-1)^{m+1-n_{ij}} z_{ij} \cup \theta_{ij}[z_j] = 0$ *be a relation in* $H^{m+1}(Z;G)$ *and* Φ *the functional cohomology operation associated with this relation. Let* $g\colon Z' \to Z$ *be a map and* Φ' *the functional cohomology operation associated with the "induced" relation* $\sum_{ij} (-1)^{m+1-n_{ij}} g^*(z_{ij}) \cup \theta_{ij}[g^* z_j] = 0$ *in* $H^{m+1}(Z';G)$. *Then, if* $f\colon X \to Z'$ *is such that* Φ'_f *is defined, we have*

$$\Phi_{gf} \subset \Phi'_f.$$

Proof. First, Φ'_f is defined if and only if $f^*(g^*z_j)=0$ in which case Φ_{gf} is also defined. Secondly, the indeterminacy of Φ'_f is

$$f^*H^m(Z')+\sum_{ij}f^*g^*z_{ij}\cup\Sigma(\theta_{ij})H^{n_j-1}(X)$$

which contains $f^*g^*H^m(Z)+\sum_{ij}f^*g^*z_{ij}\cup\Sigma(\theta_{ij})H^{n_j-1}(X)$, the indeterminacy of Φ_{gf}. To prove the theorem, therefore, we need only exhibit an element of Φ_{gf} which also belongs to Φ'_f. We will do this by looking at representative cocycles. By § 5, a representative cocycle for an element of Φ_{gf} is given by

$$v=f^\#g^\#(c)-\sum_{ij}f^\#g^\#(a_{ij})\cup\sigma(\phi_{ij})[b_j]$$

where a_{ij}, a_j, ϕ_{ij} represent z_{ij}, z_j, θ_{ij}, $f^\#g^\#(a_j)=\delta b_j$ and

$$\sum_{ij}(-1)^{m+1-n_{ij}}a_{ij}\cup\phi_{ij}[a_j]=\delta c.$$

Since $g^\#(a_{ij})$, $g^\#(a_j)$ represent $g^*z_{ij}, g^*z_j, f^\#[g^\#a_j]=\delta b_j$ and

$$\sum_{ij}(-1)^{m+1-n_{ij}}g^\#(a_{ij})\cup\phi_{ij}[g^\#a_j]=\delta g^\#c,$$

it follows that v is also a representative cocycle for an element of Φ'_f, and the theorem is proved.

9. The linearity property. The next theorem states the "linearity" property of our operations. Roughly, its content is that if one "cups" a relation on $\{z_j\}$ with a cohomology class, or if one applies to the relation a cohomology operation for which a Cartan-type formula exists, or if one takes the sum of such expressions, one still has a relation on $\{z_j\}$, and that the corresponding functional cohomology operations are related as one might expect.

To simplify the statement of the following two theorems, sufficiently complicated as they are, we assume that all coefficient groups involved coincide with a fixed ring and that the cup-products are taken with respect to the multiplication of that ring. It will be clear from the proof that the theorems remain true for any combination of coefficient groups and pairings for which the formulas make sense.

(9.1) THEOREM. *Let $z_j\in H^{n_j}(Z)$, $z_{ij,k}\in H^{m_k+1-n_{ij,k}}(Z)$, $\theta_{ij,k}$ be a cohomology operation of type $(n_j;n_{ij,k})$ and*

$$(9.2)_k \qquad \sum_{ij}(-1^{m_k+1-n_{ij,k}}z_{ij,k}\cup\theta_{ij,k}[z_j]=0$$

in $H^{m_k+1}(Z)$.

Let $\psi_{k,lm}$ be an additive cohomology operation of type $(m_k+1, m_{k,lm})$, $w_{k,lm} \in H^{n+1-m_{k,lm}}(Z)$ and suppose that

$$\psi_{k,lm}[a \cup b] = \sum_r (-1)^{i_{k,lmr}} \psi'_{k,lmr}[a] \cup \psi''_{k,lmr}[b] \tag{9.3}$$

where $i_{k,lmr}$ is the degree of $\psi'_{k,lmr}$.[2] *Then*

$$\sum_{ijklmr} (-1)^{\epsilon_{ijklmr}} w_{k,lm} \cup \psi'_{k,lmr}[z_{ij,k}] \cup \psi''_{k,lmr}\theta_{ij,k}[z_j] = 0 \tag{9.4}$$

in $H^{n+1}(Z)$ where $\epsilon_{ijklmr} = dw_{k,lm} + d\psi'_{k,lmr} + dz_{ij,k}$.

Let $\Phi^{(k)}$, Φ be the functional cohomology operations associated with $(9.2)_k$ and (9.4), respectively. If $f\colon X \to Z$ is such that $\Phi^{(k)}{}_f$ is defined, then Φ_f is also defined and

$$\Big[\sum_{klm} f^* w_{k,lm} \cup \Sigma(\psi_{k,lm})\Phi^{(k)}{}_f\Big] \cap \Phi_f \neq \phi.$$

Before we proceed to the proof of (9.1), we state a lemma which we will need.

(9.5) LEMMA. *Suppose ψ satisfies the Cartan-type formula*

$$\psi[a \cup b] = \sum_r (-1)^{i_r} \psi'_r[a] \cup \psi''_r[b]$$

where $i_r = d\psi'_r$, then $\Sigma(\psi)$, the suspension of ψ, satisfies

$$\Sigma(\psi)[a \cup b'] = \sum_r \psi'_r[a] \cup \Sigma(\psi''_r)[b']$$

where $db' = db - 1$.

For the proof, see A. Douady, [3], Exposé 9.

Proof of (9.1). Let $\operatorname{Ind}(\Phi)$ denote the indeterminacy of Φ. Then

$$\operatorname{Ind}(\Phi^{(k)}) = f^* H^{m_k}(Z) + \sum_{ij} f^* z_{ij,k} \cup \Sigma(\theta_{ij,k}) H^{n_j-1}(X),$$

$$\operatorname{Ind}(\Phi) = f^* H^n(Z) + \sum_{ijklmr} f^* w_{k,lm} \cup f^*\psi'_{k,lmr}[z_{ij,k}] \cup \Sigma(\psi''_{k,lmr}) \Sigma(\theta_{ij,k}) H^{n_j-1}(X).$$

Using (9.5), we obtain

$$\sum_{klm} f^* w_{k,lm} \cup \Sigma(\psi_{k,lm}) \operatorname{Ind}(\Phi^{(k)}) = \sum_{klm} f^*[w_{k,lm} \cup \Sigma(\psi_{k,lm}) H^{m_k}(Z)] + \sum_{ijklmr} f^* w_{k,lm} \cup \{f^*\psi'_{k,lmr}[z_{ij,k}] \cup \Sigma(\psi''_{k,lmr})\Sigma(\theta_{ij,k}) H^{n_j-1}(X)\}$$

and we have $\Big[\sum_{klm} f^* w_{k,lm} \cup \Sigma(\psi_{k,lm}) \operatorname{Ind}(\Phi^{(k)})\Big] \cap \operatorname{Ind}(\Phi) \neq \phi$.

[2] The degree, $d\psi$, of a cohomology operation ψ of type (n, q) is $q - n$.

In order to prove the theorem, we need only exhibit an element of $\sum_{klm} f^* w_{k,lm} \cup \Sigma(\psi_{k,lm}) \Phi^{(k)}{}_f$ which also belongs to Φ_f. Consider, for each k, the diagram used in § 5 to define the functional cohomology operations:

$$\begin{array}{ccccccccc} & & \sum_j H^{n_j-1}(X) & \xrightarrow{\delta} & \sum_j H^{n_j}(Z_f, X) & \xrightarrow{i^*} & \sum_j H^{n_j}(Z_f) & \xrightarrow{f'^*} & \sum_j H^{n_j}(X) \\ & & \downarrow \xi_{1,k} & & \downarrow \xi_{2,k} & & \downarrow \xi_{3,k} & & \\ H^{m_k}(Z_f) & \xrightarrow{f'^*} & H^{m_k}(X) & \xrightarrow{\delta} & H^{m_k+1}(Z_f, X) & \xrightarrow{i^*} & H^{m_k+1}(Z_f) & & \end{array}$$

where $\xi_{i,k}$ are defined as in § 5. Recall that $f'^* z'_j = 0$, so that there is z''_j with $i^* z''_j = z'_j$. Then $i^* \xi_{2,k}(\{z''_j\}) = \xi_{3,k} i^*(\{''_j\}) = 0$ and so there is $\omega_k \in H^{m_k}(X)$ such that $\delta\omega_k = \xi_{2,k}(\{z''_j\})$; $\omega_k \in \Phi^{(k)}{}_f$.

Consider now the element $\omega = \sum_{klm} f^* w_{k,lm} \cup \Sigma(\psi_{k,lm})[\omega_k]$. Let $w'_{k,lm} = r^* w_{k,lm}$. Then one verifies that

$$\delta\omega = \sum_{ijklmr} (-1)^{\epsilon_{ijklmr}} w'_{k,lm} \cup \psi'_{k,lmr}[z'_{ij,k}] \cup \psi''_{k,lmr} \theta_{ij,k}[z''_j]$$

It is now clear that $\omega \in \Phi_f$, and the theorem is proved.

10. The product theorem. The next theorem, quite similar to the preceding one, might be called the "change of variable" theorem. Basically, it states that if we have a relation on $\{z_j\}$ and then proceed to express each z_j in terms of $\{x_k\}$, other cohomology classes and cohomology operations, thus ending with a relation on $\{x_k\}$, then the corresponding two functional cohomology operations have a non-void intersection.

(10.1) THEOREM. *Let $z_j \in H^{n_j}(Z)$, $z_{ij} \in H^{m+1-n_{ij}}(Z)$, θ_{ij} be a cohomology operation of type (n_j, n_{ij}) and*

$$\sum_{ij} (-1)^{m+1-n_{ij}} z_{ij} \cup \theta_{ij}[z_j] = 0 \tag{10.2}$$

Suppose $z_j = \sum_{kl} (-1)^{dx_{jkl}} x_{jkl} \cup \psi_{jkl}[x_k]$, *$\theta_{ij}$ additive and*

$$\theta_{ij}[a \cup b] = \sum_r (-1)^{d\theta'_{ijr}} \theta'_{ijr}[a] \cup \theta''_{ijr}[b].$$

Then

$$\sum_{ijklr} (-1)^{dz_{ij}+d\theta'_{ijr}+dx_{jkl}} z_{ij} \cup \theta'_{ijr}[x_{jkl}] \cup \theta''_{ijr}\psi_{jkl}[x_k] = 0. \tag{10.3}$$

Let Φ, Ψ be the functional cohomology operations associated with (10.2), (10.3), *respectively. Then if $f\colon X \to Z$ is such that Ψ_f is defined, so is Φ_f and $\Psi_f \subset \Phi_f$.*

Proof. Let $m_k = dx_k$.

$$\operatorname{Ind}(\Phi_f) = f^*H^m(Z) + \sum_{ij} f^*z_{ij} \cup \Sigma(\theta_{ij}) H^{n_j-1}(X)$$

$$\operatorname{Ind}(\Psi_f) = f^*H^m(Z) + \sum_{ijklr} f^*z_{ij} \cup \theta'_{ijr}[f^*x_{jkl}] \cup \Sigma(\theta''_{ijr})\Sigma(\psi_{jkl}) H^{m_k-1}(X)$$

and, since

$$\begin{aligned} &\sum_r \theta'_{ijr}[f^*x_{jkl}] \cup \Sigma(\theta''_{ijr} \Sigma(\theta''_{ijr})\Sigma(\psi_{kjl}) H^{m_k-1}(X) \\ &\qquad = \Sigma(\theta_{ij})\,[f^*_{jkl} \cup \Sigma(\psi_{jkl}) H^{m_k-1}(X)] \subset \Sigma(\theta_{ij}) H^{n_j-1}(X), \end{aligned}$$

we have $\operatorname{Ind}(\Psi_f) \subset \operatorname{Ind}(\Phi_f)$.

We will now exhibit an element of Ψ_f which also belongs to Φ_f. Consider the following diagram:

$$\begin{tikzcd} \sum_k H^{m_k-1}(X) \arrow[r, "\delta"] \arrow[d, "\xi_1"] & \sum_k H^{m_k}(Z_f, X) \arrow[r, "i^*"] \arrow[d, "\xi_2"] & \sum_k H^{m_k}(Z_f) \arrow[d, "\xi_3"] \\ H^m(X) \arrow[r, "\delta"] & H^{m+1}(Z_f, X) \arrow[r, "i^*"] & H^{m+1}(Z_f) \\ \sum_k H^{n_j-1}(X) \arrow[r, "\delta"] \arrow[u, "\check{\xi}_1"] & \sum_j H^{n_j}(Z_f, X) \arrow[r, "i^*"] \arrow[u, "\bar{\xi}_2"] & \sum_j H^{n_j}(Z_f) \arrow[u, "\bar{\xi}_3"] \end{tikzcd}$$

where the top half of the diagram defines Ψ and the bottom half defines Φ, i.e.,

$$\xi_\alpha(\{x_k\}) = \sum_{ijklr} (-1)^{dz_{ij}+d\theta'_{ijr}+dx_{jkl}} z'_{ij} \cup \theta'_{ijr}[x'_{jkl}] \cup \theta''_{ijr} \cdot \phi_{jkl}[x_k],$$

$\alpha = 2, 3, \bar{\xi}_\alpha(\{z_j\}) = \sum_{ij} (-1)^{dz_{ij}} z'_{ij} \cup \theta_{ij}[z_j]$, $\alpha = 2, 3$, and ξ_1, $\bar{\xi}_1$ are defined as in § 5.

If $f^*[x_k] = 0$, so that Ψ_f is defined, then $f^*(z_j) = 0$, and Φ_f is also defined. Let $x''_k \in H^{m_k}(Z_f, X)$ satisfy $i^*(x''_k) = x'_k$, and $\omega \in \Psi_f \subset H^m(X)$, i.e., ω satisfies $\delta\omega = \xi_2(\{x''_k\})$. Letting $z''_j = \sum_{kl} (-1)^{dx_{jkl}} x'_{jkl} \cup \psi_{jkl}[x''_k]$, one verifies that $i^*(z''_j) = z'_j$ and $\delta\omega = \bar{\xi}_2(\{z''_j\})$. Thus $\omega \in \Phi_f$, and the theorem is proved.

11. Coefficients in a field. The purpose of this section is to introduce an algebraic language which seems appropriate to the study of functional cohomology operations when all cohomology classes involved have coefficients in a fixed field F.

Let A denote the Steenrod algebra over F; if X is a space, let $A(X)$

$=H^*(X;F)\otimes_F A$. We define an algebra structure on $A(X)$, quite different from the usual tensor-product of algebras, designed so that, with the resulting ring structure on $A(X)$, $H^*(X)$ becomes an $A(X)$-algebra.

Since F is a field, A is a Hopf-algebra; we denote its co-product by Δ and let $\Delta(\theta)=\sum_r \theta'_r\otimes\theta''_r$. Then we define a product in $A(X)$ by the formula:

$$(h_1\otimes\theta)\cdot(h_2\otimes\psi)=\sum_r h_1\cup\theta'_r[h_2]\otimes\theta''_r\psi.$$

Letting $A(X)$ operate on $H^*(X)$ by

$$(h\otimes\theta)(h')=h\cup\theta(h')$$

it is easily seen that

$$[(h_1\otimes\theta)\cdot(h_2\otimes\psi)](h')=(h_1\otimes\theta)[(h_2\otimes\psi)(h')]$$

so that the above definitions make $H^*(X)$ into an $A(X)$-algebra. The above construction, which can be carried out whenever a Hopf algebra acts (as a Hopf algebra) on an algebra, has been introduced independently by Massey and Peterson ([10], [11]), and called by them the *semi-tensor product* of $H^*(X)$ and A. We consider this concept, in the present paper, only for its terminological usefulness and so refer the reader to [11] for a detailed study of its properties. Our aim here is only to develop a formalism, suggested by the work of Adams [1], useful in discussing relations and associated operations.

An *X-module* is a graded free locally-finite $A(X)$-module; an *X-map* of one X-module into another is a left $A(X)$-map. A continuous map $f\colon X\to Y$ induces a map (of A-algebras) $f^*\colon A(Y)\to A(X)$, namely $f^*\otimes 1\colon H^*(Y)\otimes A\to H^*(X)\otimes A$. Let C_X, C_Y be X-, Y-modules respectively, $f\colon X\to Y$; then we say that the A-map $F\colon C_Y\to C_X$ is an *f-map* if $F(\alpha\cdot c)=f^*(\alpha)\cdot F(c)$ for all $\alpha\in A(Y)$, $c\in C_Y$.

We can now define an *X-relation* to be a pair (ϵ,x) where $\epsilon\colon C\to H^*(X)$ is a degree-preserving X-map of the X-module C into $H^*(X)$, and x is a homogeneous element of kernel (ϵ).

Thus, if C has a basis consisting of elements e_j, $\epsilon(e_j)=x_j$ and $x=\sum_j \alpha_j e_j$, where $\alpha_j=\sum_i x_{ij}\otimes\theta_{ij}$, then

$$0=\epsilon(x)=\epsilon(\sum_j\alpha_j e_j)=\sum_j\alpha_j\epsilon(e_j)=\sum_j\alpha_j x_j=\sum_{ij}x_{ij}\cup\theta_{ij}[x_j].$$

Note that if X is an Eilenberg-MacLane space $K(F,n)$, the x_j are "stable" elements and $x_{ij}=1$, our notion of X-relation reduces to the notion of relation used by Adams.

Let $f: X \to Y$, C_X and C_Y be X- and Y-modules, respectively, having the same number of generators e_i, e'_i in each dimension, $F: C_Y \to C_X$ an f-map such that $F(e'_i) = e_i$. Suppose the following diagram is commutative:

$$\begin{tikzcd} C_Y \arrow[r, "F"] \arrow[d, "\epsilon'"] & C_X \arrow[d, "\epsilon"] \\ H^*(Y) \arrow[r, "f^*"] & H^*(X) \end{tikzcd}$$

and $x = F(y)$. Then, if (ϵ', y) is a Y-relation, it is trivial that (ϵ, x) is an X-relation, *induced* from (ϵ', y) by f.

Theorems (8.1) and (9.1) may now be restated as follows:

(11.1) If Φ' is the functional cohomology operation associated with (ϵ', y), $f: X \to Y$ and Φ is the functional cohomology operation associated with the induced relation (ϵ, x), then whenever $\Phi_g(g: K \to X)$ is defined, we have

$$\Phi'_{fg} \subset \Phi_g.$$

(11.2) Let (ϵ, x_k) be X-relations, $\alpha_k \in A(X)$, $\Phi^{(k)}$ the functional cohomology operation associated with (ϵ, x_k). Then $(\epsilon, \sum_k \alpha_k x_k)$ is also an X-relation. If Φ is the functional cohomology operation associated with it, and $f: K \to X$ is such that $\Phi^{(k)}{}_f$ are defined, then

$$\sum_k f^*(\alpha_k) \cdot \Phi^{(k)}{}_f = \Phi_f$$

modulo the total indeterminacy.

It is clear from the above formulation why we have chosen to call (9.1) the "linearity" property. Rather than restate (10.1), let us state a slight generalization of it, which is an immediate consequence of (8.1) and (10.1).

(11.3) Let $f: X \to Y$, C_X and C_Y be X- and Y-modules, respectively, $F: C_Y \to C_X$ an f-map. Suppose the following diagram commutative:

$$\begin{tikzcd} C_Y \arrow[r, "F"] \arrow[d, "\epsilon'"'] & C_X \arrow[d, "\epsilon"] \\ H^*(Y) \arrow[r, "f^*"] & H^*(X) \end{tikzcd}$$

and $x = F(y)$. Then if Φ is the functional cohomology operation associated

with (ϵ', y) and Ψ that associated with (ϵ, x), we have, whenever $g: K \to X$ is such that Ψ_g is defined,

$$\Psi_g = \Phi_{fg}$$

modulo the total indeterminacy.

Chapter V. Applications.

We will present in this chapter certain applications of the results of preceding chapters. These will be of two kinds: (1) we will complete the results of Peterson-Stein [18] and give formulas of the Wu- and Whitney-type for iterated Steenrod squares, $\mathrm{Sq}^I\Phi_J$, of secondary characteristic classes of sphere-bundles and for the secondary characteristic classes $\Phi_J(\xi \oplus \xi')$ of the Whitney sum of two such bundles, (2) we will study Moore-Postnikov systems of fibration and show that, under suitable "stability" conditions, each k-invariant belongs to a functional cohomology operation associated with a certain relation on the preceding k-invariant.

12. Secondary characteristic classes. Let G_n be the oriented Grassman manifold of oriented n-planes through the origin in R_∞. Let $W_i \in H^i(G_n : Z_2)$, $i = 2, \cdots, n$, be the universal Stiefel-Whitney classes and $X \in H^n(G_n : Z)$ the universal Euler class. Consider the natural set of mappings T into G_n defined by the Euler class X. This set coincides with the set of classifying maps of oriented $(n-1)$-sphere bundles with vanishing Euler class. The Wu formulas for the Steenrod squares of W_i yield relations in G_n of the form

$$\mathrm{Sq}^I W_n + Q_I \cup W_n = 0 \tag{12.1}$$

where Q_I is a polynomial in $W_2, \cdots, W_n$, $I = \{i_1 \cdots i_r\}$, $i_1, \cdots, i_r$ non-negative integers, $\mathrm{Sq}^I = \mathrm{Sq}^{i_1} \cdots \mathrm{Sq}^{i_r}$. Let $n(I) = \sum_{j=1}^{r} i_j$. Let Φ^I be the functional cohomology operation associated with (12.1). Then if $f(\xi)$: $B(\xi) \to G_n$ is the classifying map of an oriented $(n-1)$-sphere bundle ξ with vanishing Euler class $X(\xi)$ (and therefore with vanishing $W_n(\xi)$), the secondary characteristic classes defined by Peterson-Stein [18] are closely connected with $\Phi^I{}_{f(\xi)}$. Let $\pi: K_n \to G_n$ be the universal example for T; then $\Phi^I{}_\pi$ is defined, is a coset of $\pi^* H^{n-1+n(I)}(G_n) + [\mathrm{Sq}^I + \pi^*(Q_I) \cup] H^{n-1}(K_n)$ and is the universal example for Φ^I. Peterson and Stein show that one can single out a subset $\Phi_{I,\pi}$ of $\Phi^I{}_\pi$ which is a coset of $[\mathrm{Sq}^I + \pi^*(Q_I) \cup] H^{n-1}(K_n)$ and which can then be taken as the universal example for a "smaller" functional cohomology operation Φ_I. Their secondary characteristic classes are $\Phi_{I,f(\xi)}$ and are denoted by $\Phi_I(\xi)$; similarly let us denote $\Phi^I{}_{f(\xi)}$ by $\Phi^I(\xi)$.

Clearly, for any ξ, $\Phi_I(\xi) \subset \Phi^I(\xi)$ and any element of $\Phi^I(\xi)$ may be written as a suitable element of $\Phi_I(\xi)$ plus a polynomial in the Stiefel-Whitney classes $W_i(\xi)$. Thus formulas involving the $\Phi^I(\xi)$ imply corresponding formulas for the $\Phi_I(\xi)$ only modulo primary characteristic classes and are therefore correspondingly weaker. Our formulas will deal with the Φ^I; it seems likely that analogous formulas for the Φ_I can be obtained by a more careful analysis, but we shall not bother to do this here.

We will need some lemmas. If $I' = (i'_1, \cdots, i'_r)$, $I'' = (i''_1, \cdots, i''_r)$, then $I' + I'' = (i'_1 + i''_1, \cdots, i'_r + ''_r)$.

(12.2) LEMMA.

$$\mathrm{Sq}^I(x \cup y) = \sum_{I'+I''=I} \mathrm{Sq}^{I'}(x) \cup \mathrm{Sq}^{I''}(y).$$

The proof is by induction on r, the length of I, using the Cartan formula.

(12.3) LEMMA.

$$Q_{J,I} = \sum_{J'+J''=J} \mathrm{Sq}^{J'} Q_I \cup Q_{J''}.$$

Proof.

$$\begin{aligned} Q_{J,I} \cup W_n &= \mathrm{Sq}^{(J,I)} W_n = \mathrm{Sq}^J \mathrm{Sq}^I W_n \\ &= \mathrm{Sq}^J (Q_I \cup W_n) \\ &= \sum_{J'+J''=J} \mathrm{Sq}^{J'} Q_I \cup \mathrm{Sq}^{J''} W_n \\ &= \sum_{J'+J''=J} \mathrm{Sq}^{J'} Q_I \cup Q_{J''} \cup W_n. \end{aligned}$$

Since $H^*(G_n; Z_2)$ is a polynomial algebra, the result follows.

(12.4) THEOREM.

$$\mathrm{Sq}^J \Phi^I(\xi) = \Phi^{(J,I)}(\xi) + \sum_{J'+J''=J} \mathrm{Sq}^{J'} Q_I(\xi) \cup \Phi^{J''}(\xi)$$

modulo the total indeterminacy which is

$$\begin{aligned} f(\xi)^* H^{n+n(I)+n(J)-1}(G_n) &+ \sum \mathrm{Sq}^{J'} Q_I \cup (\mathrm{Sq}^{J''} + Q_{J''} \cup) H^{n-1}(B(\xi)) \\ &+ [\sum (\mathrm{Sq}^{J'} Q_I \cup Q_{J''} \cup) + \mathrm{Sq}^J \mathrm{Sq}^I] H^{n-1}(B(\xi)). \end{aligned}$$

Proof. Φ^I is associated with the relation $\mathrm{Sq}^I W_n + Q_I \cup W_n = 0$. Now,

$$\mathrm{Sq}^J(\mathrm{Sq}^I W_n + Q_I \cup W_n) = \mathrm{Sq}^J \mathrm{Sq}^I W_n + \sum_{J'+J''=J} \mathrm{Sq}^{J'} Q_I \cup \mathrm{Sq}^{J''} W_n,$$

and so by theorem (9.1), $\mathrm{Sq}^J \Phi^I(\xi)$ intersects $\Psi_{f(\xi)}$ where Ψ is the functional cohomology operation associated with the relation

$$\mathrm{Sq}^J \mathrm{Sq}^I W_n + \sum \mathrm{Sq}^{J'} Q_I \cup \mathrm{Sq}^{J''} W_n = 0.$$

On the other hand,

$$\begin{aligned}\mathrm{Sq}^{(J,I)}W_n + \sum \mathrm{Sq}^{J'}Q_I \cup \mathrm{Sq}^{J''}W_n &= \mathrm{Sq}^{(J,I)}W_n + Q_{(J,I)} \cup W_n \\ &\quad + Q_{(J,I)} \cup W_n + \sum \mathrm{Sq}^{J'}Q_I \cup \mathrm{Sq}^{J''}W_n \\ &= [\mathrm{Sq}^{(J,I)}W_n + Q_{(J,I)} \cup W_n] + \sum \mathrm{Sq}^{J'}Q_I \cup (\mathrm{Sq}^{J''}W_n + Q_{J''} \cup W_n)\end{aligned}$$

by Lemma (12.3), and so, by another application of Theorem (9.1), we obtain

$$\Phi^{(J,I)}(\xi) + \sum \mathrm{Sq}^{J'}Q_I(\xi) \cup \Phi^{J''}(\xi) = \Psi_{f(\xi)}$$

modulo the total indeterminacy. The result now follows upon computing the various indeterminacies.

This result is to be compared with Theorems (6.2) and (6.3) of [18]. Comparison with (6.3) will illustrate the fact that every Adem relation between Sq^I's yields a similar relation between the corresponding Φ^I's. These are proved, of course, using Theorem (9.1).

Let now ξ be an oriented $(m-1)$-sphere bundle over B, ξ' an oriented $(n-1)$-sphere bundle over the same base B. Suppose $X(\xi)=0$. Then $X(\xi\oplus\xi')=0$. We wish to compute $\Phi^I(\xi\oplus\xi')$. Let $f(\xi): B\to G_m$, $f(\xi'): B\to G_n$ be classifying maps for ξ, ξ', respectively. It is well-known that the classifying map for $\xi\oplus\xi'$ may be described as follows: let $\mu\colon G_m\times G_n\to G_{m+n}$ be the map which associates to a pair of m- and n-planes in R_∞ the $m+n$-plane spanned by them. Then the composition

$$B \xrightarrow{\Delta} B\times B \xrightarrow{f(\xi)\times f(\xi')} G_m\times G_n \xrightarrow{\mu} G_{m+n}$$

is $f(\xi\oplus\xi')$. Let ω_i, W'_i, W''_i denote the universal Stiefel-Whitney classes in G_{m+n}, G_m, and G_n, respectively and similarly for $\bar{Q}_I$, Q'_I, Q''_I. The Whitney duality theorem states that $\mu^*\omega_i = \sum_{\alpha+\beta=i} W'_\alpha\otimes W''_\beta$.

(12.5) LEMMA.

$$\mu^*\bar{Q}_I = \sum_{I'+I''=I} Q'_{I'}\otimes Q''_{I''}.$$

Proof. Since $\bar{Q}_I\cup\omega_{m+n} = \mathrm{Sq}^I\omega_{m+n}$, we have

$$\mu^*(\bar{Q}_I\cup\omega_{m+n}) = \mu^*\bar{Q}_I\cup\mu^*\omega_{m+n} = \mu^*\bar{Q}_I\cup(W'_m\otimes W''_n),$$

and

$$\begin{aligned}\mu^*(\bar{Q}_I\cup\omega_{m+n}) &= \mu^*(\mathrm{Sq}^I\omega_{m+n}) = \mathrm{Sq}^I\mu^*\omega_{m+n} = \mathrm{Sq}^I(W'_m\otimes W''_n) \\ &= \sum \mathrm{Sq}^{I'}W'_m\otimes \mathrm{Sq}^{I''}W''_n = \sum Q'_{I'}W'_m\otimes Q''_{I''}W''_n \\ &= (\sum Q'_{I'}\otimes Q''_{I''})\cup(W'_m\otimes W''_n).\end{aligned}$$

Since $H^*(G_m\times G_n; Z_2) = Z_2[W'_2,\cdots,W'_m; W''_2,\cdots,W''_n]$, the result follows.

(12.6) THEOREM.

$$\Phi^I(\xi\oplus\xi') = \sum_{I'+I''=I}\Phi^{I'}(\xi)\cup Q_{I''}(\xi')\cup W_n(\xi')$$

modulo the total indeterminacy which is

$$f(\xi\oplus\xi')^*H^{m+n-1+n(I)}(G_{m+n})+[\mathrm{Sq}^I+\Sigma\, Q_{I'}(\xi)Q_{I''}(\xi')\cup]H^{m+n-1}(B)$$
$$+\Sigma\,[\mathrm{Sq}^{I'}+Q_{I'}(\xi)\cup]H^{m-1}(B)\cup Q_{I''}(\xi')W_n(\xi')$$

Proof. Let $\tilde{\Phi}^I$ be the functional cohomology operation associated with the relation $\mathrm{Sq}^I\omega_{m+n}+\tilde{Q}_I\cup\omega_{m+n}=0$ in G_{m+n}. Then

$$\Phi^I(\xi\oplus\xi')=\tilde{\Phi}^I{}_{f(\xi\oplus\xi')}=\tilde{\Phi}^I{}_{\mu(f(\xi)\times f(\xi'))\Delta}.$$

Now,

$$\begin{aligned}\mu^*(\mathrm{Sq}^I\omega_{m+n}+\tilde{Q}_I\cup\omega_{m+n})&=\mathrm{Sq}^I\mu^*\omega_{m+n}+\mu^*(\tilde{Q}_I\cup\omega_{m+n})\\&=\mathrm{Sq}^I(W'_m\otimes W''_n)+\mu^*(\tilde{Q}_I)\cup(W'_m\otimes W''_n)\\&=\sum_{I'+I''=I}\mathrm{Sq}^{I'}W'_m\otimes\mathrm{Sq}^{I''}W''_n+\sum_{I'+I''=I}(Q'_{I'}\otimes Q''_{I''})\cup(W'_m\otimes W''_n)\\&=\sum_{I'+I''=I}(\mathrm{Sq}^{I'}W'_m\otimes\mathrm{Sq}^{I''}W''_n+Q'_{I'}W'_m\otimes Q''_{I''}W''_n)\\&=\sum_{I'+I''=I}(\mathrm{Sq}^{I'}W'_m+Q'_{I'}W'_m)\otimes Q''_{I''}W''_n.\end{aligned}$$

Let Ψ, Ψ', Φ'^I be the functional cohomology operations associated with the following relations, respectively:

$$\Psi:\ \mathrm{Sq}^I(W'_m\otimes W''_n)+\mu^*(\tilde{Q}_I)\cup(W'_m\otimes W''_n)=0\qquad(\text{variable}=W'_m\otimes W''_n)$$

$$\Psi':\ \sum_{I'+I''=I}(\mathrm{Sq}^{I'}W'_m+Q'_{I'}W'_m)\otimes Q''_{I''}W''_n=0\qquad(\text{variable}=W'_m)$$

$$\Phi'^I:\ (\mathrm{Sq}^IW'_m+Q'_IW'_m)\otimes 1=0\qquad(\text{variable}=W'_m)$$

Then the following hold:

$$(12.7)\quad\begin{aligned}&\tilde{\Phi}^I{}_{\mu(f(\xi)\times f(\xi'))\Delta}\subset\Psi_{(f(\xi)\times f(\xi'))\Delta}\ \text{by (8.1)}\\&\Psi'_{(f(\xi)\times f(\xi'))\Delta}\subset\Psi_{(f(\xi)\times f(\xi'))\Delta}\ \text{by (10.1)}\\&[\sum_{I'+I''=I}\Phi'^{I'}{}_{(f(\xi)\times f(\xi'))\Delta}\cup[(f(\xi)\times f(\xi'))\Delta]^*Q''_{I''}W''_n]\\&\qquad\cap\Psi'_{(f(\xi)\times f(\xi'))\Delta}\neq\phi\ \text{by (9.1)}.\end{aligned}$$

Now, $[(f(\xi)\times f(\xi'))\Delta]^*Q''_{I''}W''_n=Q_{I''}(\xi')W_n(\xi')$; finally, since

$$(\mathrm{Sq}^IW'_m+Q'_I\cup W'_m)\otimes 1=0$$

is induced from $\mathrm{Sq}^IW'_m+Q'_I\cup W'_m=0$ by the projection $\pi_1\colon G_m\times G_n\to G_m$, and $\pi_1(f(\xi)\times f(\xi'))\Delta=f(\xi)$, we have

$$(12.8)\qquad\Phi'^{I'}{}_{(f(\xi)\times f(\xi'))\Delta}\supset\Phi^{I'}{}_{f(\xi)}.$$

Combining (12.7) and (12.8) proves our theorem.

13. Moore-Postnikov systems. In this section, we prove a number

of results on the Moore-Postnikov systems of fibrations. First, let us recall some definitions and terminology [6], [7].

Let $F \xrightarrow{i} E \xrightarrow{p} B$ be a fibration such that $\pi_1(B)$ operates trivially on the homology and homotopy of the fiber; let $0 < n(1) < n(2) < \cdots$ be the dimensions in which F has non-trivial homotopy groups, $\pi_j = \pi_{n(j)}(F)$. A Moore-Postnikov system of this fibration is a sequence of fibrations $K(\pi_\alpha, n(\alpha)) \xrightarrow{i_\alpha} E_\alpha \xrightarrow{p_\alpha} E_{\alpha-1}$, maps $q_\alpha \colon E \to E_\alpha$ such that $E_0 = B$, $q_0 = p$, $q_{\alpha-1} = p_\alpha q_\alpha$, and q_α induces isomorphisms of homotopy groups in dimensions $\leqq n(\alpha)$. A Postnikov system of a space F is a Moore-Postnikov system of the fibration $F \to F \to$ point. "Restricting" the Moore-Postnikov system of p to that portion of it "lying above the basepoint $*$ of B" yields, as is easily seen, a Postnikov system of F, $K(\pi_\alpha, n(\alpha)) \xrightarrow{j_\alpha} F_\alpha \xrightarrow{r_\alpha} F_{\alpha-1}$ such that the diagram below is commutative:

$$\begin{tikzcd}
 & E_\alpha \arrow[r, "p_\alpha"] & E_{\alpha-1} \\
K(\pi_\alpha, n(\alpha)) \arrow[ur, "i_\alpha"] \arrow[dr, "j_\alpha"'] & & \\
 & F_\alpha \arrow[r, "r_\alpha"] \arrow[uu, "l_\alpha"] & F_{\alpha-1} \arrow[uu, "l_{\alpha-1}"']
\end{tikzcd}$$

where l_α, $l_{\alpha-1}$ are the inclusions. Let $b^\alpha \in H^{n(\alpha)}(K(\pi_\alpha, n(\alpha)); \pi_\alpha)$ be the basic class, then the α-th k-invariants of p, F, respectively are defined by $k^\alpha = -\tau_p(b^\alpha)$, $k^\alpha(F) = -\tau_r(b^\alpha)$ where τ_p, τ_r are the transgressions. Thus $k^\alpha \in H^{n(\alpha)+1}(E_{\alpha-1}; \pi_\alpha)$, $k^\alpha(F) \in H^{n(\alpha)+1}(F_{\alpha-1}; \pi_\alpha)$, and $l_{\alpha-1}{}^* k^\alpha = k^\alpha(F)$. The element $i_{\alpha-1}{}^* k^\alpha - j_{\alpha-1}{}^* k^\alpha(F)$ of $H^{n(\alpha)+1}(\pi_{\alpha-1}, n(\alpha-1); \pi_\alpha)$ determines a cohomology operation of type $(n(\alpha-1), n(\alpha)+1; \pi_{\alpha-1}, \pi_\alpha)$, called the α-th *primary* k-invariant of F, of of p, and denoted by θ^α. We will see in the next theorem that these rather crude invariants often yield a considerable amount of information concerning the k^α and $k^\alpha(F)$. If $n(\alpha+1)+1 < 2n(\alpha)$, then $\theta^{\alpha+1}$ can be "desuspended" uniquely, i. e., there is a unique cohomology operation $\Psi^{\alpha+1}$ of type $(n(\alpha)+1, n(\alpha+1)+2; \pi_\alpha, \pi_{\alpha+1})$ such that $\Sigma(\Psi^{\alpha+1}) = \theta^{\alpha+1}$, [5], I.

(13.1) THEOREM. *Suppose $n(\alpha+1)+1 < 2n(\alpha)$ and $\pi_{\alpha+1}$ is finitely-generated, then there is a uniquely defined relation of the form*

$$\Psi^{\alpha+1}[k^\alpha] + \sum_\lambda h_\lambda \cup \Psi_\lambda[k^\alpha] = 0 \qquad (dh_\lambda > 0) \tag{13.2}$$

in $H^{n(\alpha+1)+2}(E_{\alpha-1}; \pi_{\alpha+1})$ such that $k^{\alpha+1} \in \Phi_{p_\alpha}$, where Φ is the functional cohomology operation associated with (13.2).

Proof. Let $\pi_{\alpha+1} = \sum_s \Delta_s$ be the canonical decomposition of $\pi_{\alpha+1}$ into the

direct sum of cyclic groups such that the order of any finite summand divides the order of the following finite summands. Thus each Λ_s is a ring, and we can apply (7.1) to $H^{n(\alpha+1)+1}(E_\alpha;\Lambda_s)$. Let now

$$k^{\alpha+1}=\sum_s k_s^{\alpha+1}\in\sum_s H^{n(\alpha+1)+1}(E_\alpha;\Lambda_s)=H^{n(\alpha+1)+1}(E_\alpha;\pi_{\alpha+1}).$$

By (7.1), since $n(\alpha+1)+1<2n(\alpha)$, there is a unique relation of the form

$$\sum_\eta h_{s,\eta}\cup\Psi_{s,\eta}[k^\alpha]=0 \tag{13.3}$$

such that $k_s^{\alpha+1}\in\Phi^{(s)}{}_{p_\alpha}$ where $\Phi^{(s)}$ is the functional cohomology operation associated with (13.3). In (13.3), $\Psi_{s,\eta}$ denotes a (primary) cohomology operation of type (π_α,Λ_s). By Theorem (9.1), we see that $k^{\alpha+1}=\sum_s k_s^{\alpha+1}$ $\in\sum_s\Phi^{(s)}{}_{p_\alpha}=\Phi_{p_\alpha}$ where Φ is the functional cohomology operation associated with the relation $\sum_{s\eta}h_{s,\eta}\cup\Psi_{s,\eta}[k^\alpha]=0$. It still remains to verify that in this last relation, the terms with $dh_{s,\eta}=0$ add up to $\Psi^{\alpha+1}[k^\alpha]$. To see this, recall that $i_\alpha{}^*k^{\alpha+1}=\Sigma(\Psi^{\alpha+1})[b^\alpha]$ and that

$$\tau\Sigma(\Psi^{\alpha+1})[b^\alpha]=\Psi^{\alpha+1}\tau(b^\alpha)=-\Psi^{\alpha+1}[k^\alpha];$$

the result now follows from Theorem (7.4).

(13.4) COROLLARY. *If F is $(n-1)$-connected, then the conclusion of* (13.1) *holds for all $k^{\alpha+1}$ such that $n(\alpha+1)<2n-1$.*

Let us recall from [13] the definition of an H-fibration: $F\xrightarrow{i}E\xrightarrow{p}B$ is an H-fibration if there is a map $\mu\colon F\times E\to E$ and a homotopy $H_t\colon F\vee E\to E$ such that $p\mu=p\pi_2$, $H_0=\mu j$, $H_1=\nabla(i\vee 1)$, $pH_t(F\vee F)=*$, where $j\colon F\vee E\subset F\times E$, $\nabla\colon E\vee E\to E$ is the folding map. F is then an H-space; one proves, as in the case of an H-space, that the k-invariants of p are primitive.

(13.5) COROLLARY. *Let $F\xrightarrow{i}E\xrightarrow{p}B$ be an H-fibration and $n(\alpha+1)+1<2n(\alpha)$, then $\Psi^{\alpha+1}[k^\alpha]=0$.*

Proof. Let μ also denote the operation $K(\pi_\alpha,n(\alpha))\times E_\alpha\to E_\alpha$. Then $\mu^*k^{\alpha+1}=\pi_2{}^*k^{\alpha+1}+\pi_1{}^*i_\alpha{}^*k^{\alpha+1}$, $\mu^*k_s{}^{\alpha+1}=\pi_2{}^*k_s{}^{\alpha+1}+\pi_1{}^*i_\alpha{}^*k_s{}^{\alpha+1}$, for all s, and so by (7.4), we must have $\sum_\eta'\pm\pi_2{}^*p_\alpha{}^*h_{s,\eta}\cup\pi_1{}^*\theta_{s,\eta}[b^\alpha]=0$ where the sum $\sum_\eta'$ is taken over all η such that $d(h_{s,\eta})>0$. Hence

$$\sum_\eta'\pm\theta_{s,\eta}[b^\alpha]\times p_\alpha{}^*h_{s,\eta}=0,$$

and $\theta_{s,\eta}[b^\alpha]=0$ or $p_\alpha{}^*h_{s,\eta}=0$. Now

$$dh_{s,\eta}\leqq n(\alpha+1)+2-n(\alpha)-1=n(\alpha+1)-n(\alpha)+1<n(\alpha)$$

and p_α^* is an isomorphism in dimensions $< n(\alpha)$; on the other hand, $\theta_{s,\eta}$ determines uniquely $\psi_{s,\eta}$ with $\Sigma(\psi_{s,\eta}) = \theta_{s,\eta}$, and we obtain $h_{s,\eta} = 0$ or $\psi_{s,\eta} = 0$ for every (s,η) such that $dh_{s,\eta} > 0$.

Let us now apply the preceding to the study of Postnikov systems of spaces. Thus $B = *$ and $F = E$.

(13.6) COROLLARY. *A necessary condition for a space E to be an H-space is that $\Psi^{\alpha+1}[k^\alpha] = 0$ whenever $n(\alpha+1)+1 < 2n(\alpha)$.*

(13.7) COROLLARY.

(1) *If $n(\alpha+1) - n(\alpha) + 1 < n(1)$, then $\Psi^{\alpha+1}[k^\alpha] = 0$.*

(2) *If $n(\alpha+1) - n(\alpha) + 1 = n(1)$, then*
$$\Psi^{\alpha+1}[k^\alpha] + (-1)^{n(1)} b^1 \cup k^\alpha = 0,$$
where the cup-product is taken with respect to the Whitehead product $W\colon \pi_1 \otimes \pi_\alpha \to \pi_{\alpha+1}$.

Proof. In (13.2), $dh_\lambda \leqq n(\alpha+1)+2-n(\alpha)-1 = n(\alpha+1)-n(\alpha)+1$. Thus, in case (1), $h_\lambda = 0$ in positive dimensions, and (1) is proved. In case (2), $dh_\lambda = 0$ or $n(1)$ and so $h_\lambda = 0$ or $a_\#[b^1]$, where $a_\#$ is the coefficient homomorphism induced from $a\colon \pi_1 \to G$, and b^1 is the basic class of $E_{\alpha-1}$; $d\psi_\lambda = 0$ and $\psi_\lambda = a'_\#$, where $a'\colon \pi_\alpha \to G'$ so that

$$\Psi^{\alpha+1}[k^\alpha] + a_\#[b^1] \cup a'_\#[k^\alpha] = 0$$

where the cup-product is taken with respect to a pairing $b\colon G \times G' \to \pi_{\alpha+1}$. By (7.4) we have

$$\mu^* k^{\alpha+1} = \pi_2^* k^{\alpha+1} + \pi_1^* i_\alpha^* k^{\alpha+1} + (-1)^{n(1)} \pi_2^* p_\alpha^* a_\#[b^1] \cup \pi_1^* a'_\#[b^\alpha],$$

where $\mu\colon K(\pi_\alpha, n(\alpha)) \times E_\alpha \to E_\alpha$. It follows from Theorem (3.3) of [12] that the diagram below is commutative:

$$\begin{tikzcd}
\pi_1 \otimes \pi_\alpha \arrow[r, "a \otimes a'"] \arrow[dr, "(-1)^{n(1)} W"'] & G \otimes G' \arrow[d, "b"] \\
 & \pi_{\alpha+1}
\end{tikzcd}$$

and the corollary is now proved.

(13.8) *Remark.* In view of the above, it is clear that a good knowledge of the Ψ^α would be quite helpful. It is classical that, if $n(\alpha) = n(\alpha-1)+1$, then Ψ^α is Sq^2 with respect to the homomorphism $\eta\colon \pi_{\alpha-1} \to \pi_\alpha$ induced by composition with the non-trivial element of $\pi_{n(\alpha)}(S^{n(\alpha-1)})$. The only other results in this direction are due to Nakaoka [14], [15].

14. Further applications and remarks.

(14.1) Adem (Sobre la formula del producto para operaciones cohomologicas de segundo orden, Bol. Soc. Mat. Mexicana, 4 (1959)) and Kristensen (On secondary cohomology operations, Mathematica Scandinavica, 12 (1963)) define and study secondary cohomology operations via the use of cochain formulas. Their formulas are very similar to ours. They are more explicit for they use only Steenrod's cup-i-products instead of arbitrary cochain operations. Many complicated computations of theirs can be avoided by using the theory of cochain operations.

(14.2) Adem, in a recent paper (Sobre operaciones cohomologicas secundarias, Bol. Soc. Mat. Mexicana, 7 (1962)) continues his earlier work; the emphasis now, as in the present paper, is on the "diagram" approach and he proves a number of product formulas for secondary and functional cohomology operations. These may be considered as easy consequences of Theorems (8.1), (9.1), and (10.1) of the present paper. The arguments would be very similar to, although easier than those of § 12. This point of view, incidentally, shows that even if one is only interested in operations arising from relations in the Steenrod algebra, the question of product formulas for such operations naturally leads to the study of the more complicated relations discussed here.

(14.3) W. C. Hsiang, in another recent paper (Higher obstructions to sectioning a special type of fibre bundles, Trans. Amer. Math. Soc., 110 (1964)) develops the theory of "enlargement" to deal with certain problems of obstruction theory. One of his applications gives information on the third obstruction in (CP^{n-1}, PU)-bundles. We wish to sketch here how his result may be obtained by the "orthodox" method of Postnikov systems, using the results of § 13.

Let $CP^{n-1} \to E \xrightarrow{P} B$ be a fibration, with CP^{n-1} the complex projective space of $2n-2$ real dimensions, and with vanishing first obstruction. In this situation, Hermann [7] exhibits classes $a(f) \in H^2(E;Z)$, $b_j \in H^{2n-2j}(B;Z)$ such that $a(f)^n = \sum_{j=0}^{n-1} p^*(b_j) \cup a(f)^j$. Consider the Moore-Postnikov system of $p\colon E_1 = B \times K(Z,2)$, $k^2 \in H^{2n}(E_1;Z)$, $k^3 \in H^{2n+1}(E_2;Z_2)$. Hermann shows that $k^2 = 1 \times \iota^n - \sum_{j=0}^{n-1} b_j \times \iota^j$, where ι is the basic class of $K(Z,2)$; by § 13, k^3 arises out of a unique relation of the form $\mathrm{Sq}^2 k^2 + \cdots = 0$. This relation is easily computed using the above information; the result is $\mathrm{Sq}^2 k^2 + [b_{n-1} \times 1 + n(1 \times \iota)] \cup k^2 = 0$. In the special case where E is a

fibre-bundle with group PU, comparison of Kundert's classical work and Hermann's shows that there is an associated S^{2n-1}-bundle with group U whose Chern classes c_i satisfy $c_{n-j} = -b_j$. Hsiang's Theorem (6.1) now follows since $b_{n-1} = -c_1 \equiv W_2 \pmod 2$.

REFERENCES.

[1] J. F. Adams, "On the non-existence of elements of Hopf invariant one," *Annals of Mathematics* (2), vol. 72 (1960), pp. 20-104.

[2] H. Cartan, "Algèbres d'Eilenberg-MacLane et homotopie," *Séminaire Henri Cartan 1954-1955*, Paris.

[3] ———, "Invariant de Hopf et opérations cohomologiques secondaires," *Séminaire Henri Cartan 1958-1959*, Paris.

[4] W. H. Cockcroft, "The cohomology groups of a fibre space with fibre a space of type $K(\pi, n)$, II," *Transactions of the American Mathematical Society*, vol. 91 (1959), pp. 505-524.

[5] S. Eilenberg and S. MacLane, "On the groups $H(\pi, n)$, I," *Annals of Mathematics* (2), vol. 58 (1953), pp. 55-106; II, *ibid.*, vol. 60 (1954), pp. 49-139; III, *ibid.*, vol. 60 (1954), pp. 513-557.

[6] R. Hermann, "Secondary obstructions for fibre spaces," *Bulletin of the American Mathematical Society*, vol. 65 (1959), pp. 5-8.

[7] ———, "Obstruction theory for fibre spaces," *Illinois Journal of Mathematics*, vol. 4 (1960), pp. 9-27.

[8] G. Hirsch, "Sur les groupes d'homologie des espaces fibrés," *Bulletin de la Société Mathématique de Belgique*, vol. 6 (1953), pp. 79-96.

[9] M. Mahowald, *The third and fourth obstructions in sphere-bundles*, mimeographed notes.

[10] W. S. Massey and F. P. Peterson, "On the cohomology algebra of fibre-bundles whose fibre is totally non-homologous to zero," *Reports of Seattle conference in differential topology*, Summer 1963, mimeographed.

[11] ———, detailed version of [10], *to appear*.

[12] J.-P. Meyer, "Whitehead products and Postnikov systems," *American Journal of Mathematics*, vol. 81 (1960), pp. 271-280.

[13] ———, "Principal fibrations," *Transactions of the American Mathematical Society*, vol. 107 (1963), pp. 177-185.

[14] M. Nakaoka, "Classification of mappings of a complex into a special kind of complex," *J. Inst. Poly., Osaka City Univ.*, vol. 3 (1952), pp. 101-143.

[15] ———, "On homotopy classification and extension," *Proceedings of the Japan Academy*, vol. 29 (1953), pp. 6-9.

[16] F. P. Peterson, "Functional cohomology operations," *Transactions of the American Mathematical Society*, vol. 86 (1957), pp. 187-197.

[17] ———, "Whitehead products and the cohomology structure of principal fibre spaces," *American Journal of Mathematics*, vol. 82 (1960), pp. 649-652.

[18] F. P. Peterson and N. Stein, "Secondary characteristic classes," *Annals of Mathematics* (2), vol. 76 (1962), pp. 510-523.

[19] F. P. Peterson and E. Thomas, "A note on non-stable cohomology operations," *Boletín de la Sociedad Matemática Mexicana* (1958), pp. 13-18.

[20] N. E. Steenrod, "Cohomology invariants of mappings," *Annals of Mathematics* (2), vol. 50 (1949), pp. 954-988.

A COMBINATORIAL PROBLEM CONNECTED WITH DIFFERENTIAL EQUATIONS.

By H. Davenport and A. Schinzel.

1. Let

$$F(D)f(x)=0 \tag{1}$$

be a (homogeneous) linear differential equation with constant coefficients, of order d. Suppose that $F(D)$ has real coefficients, and that the roots of $F(\lambda)=0$ are all real though not necessarily distinct. As is well known, any solution of (1) is of the form

$$f(x)=P_1(x)e^{\lambda_1 x}+\cdots+P_k(x)e^{\lambda_k x}, \tag{2}$$

where $\lambda_1,\cdots,\lambda_k$ are the distinct roots of $F(\lambda)=0$ and $P_1(x),\cdots,P_k(x)$ are polynomials of degrees at most $m_1-1,\cdots,m_k-1$, where $m_1,\cdots,m_k$ are the multiplicities of the roots, so that $m_1+\cdots+m_k=d$.

Let

$$f_1(x),\cdots,f_n(x) \tag{3}$$

be n distinct (but not necessarily independent) solutions of (1). For each real number x, apart from a finite number of exceptions, there will be just one of the functions (3) which is greater than all the others. We can therefore dissect the real line into N intervals

$$(-\infty,x_1),\ (x_1,x_2),\cdots,\ (x_{N-1},\infty)$$

such that inside any one of the intervals (x_{j-1},x_j) a particular one of the functions (3) is the greatest, and such that this function is not the same for two consecutive intervals. It is almost obvious that N is finite, and a formal proof will be given below.

The problem of finding how large N can be, for given d and given n, was proposed to one of us (in a slightly different form) by K. Malanowski. This problem can be made to depend on a purely combinatorial problem, by the following considerations. With each $j=1,2,\cdots,N$ there is associated the integer $i=i(j)$ for which $f_i(x)$ is the greatest of the functions (3) in the interval (x_{j-1},x_j). (We write $x_0=-\infty$ and $x_N=\infty$ for convenience.) This defines a sequence of N terms

$$i(1),i(2),\cdots,i(N), \tag{4}$$

Received August 26, 1964.

each term being one of $1, 2, \cdots, n$. This sequence has no two consecutive terms equal, which we may express by saying that it has no immediate repetition. The sequence has the further property that it contains no subsequence of the form

$$a, b, a, b, \cdots \quad \text{with } d+1 \text{ terms and } a \neq b. \tag{5}$$

For suppose that $j_1 < j_2 < \cdots < j_{d+1}$ and that

$$i(j_1) = a,\ i(j_2) = b,\ i(j_3) = a, \cdots.$$

Then the function $f_a(x) - f_b(x)$ is positive in (x_{j_1-1}, x_{j_1}), negative in (x_{j_2-1}, x_{j_2}), and so on. Hence this function has a zero between x_{j_1} and x_{j_2-1}, another zero between x_{j_2} and x_{j_3-1}, and so on, making at least d distinct zeros. But $f_a(x) - f_b(x)$ is itself a function of the type (2), and it is known [1] that any such function has at most $d-1$ zeros.

[1] See, for example, G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, II, Section V, Problem 75.

We are therefore led to the following combinatorial problem: *to find the greatest length of a sequence with no immediate repetition, each term of which is one of* $1, 2, \cdots, n$, *and which contains no subsequence of the type* (5). We shall denote this greatest length (that is, greatest number of terms) by $N_d(n)$. Any upper bound obtained for $N_d(n)$ will be valid for the number N defined earlier in relation to the differential equation (1). We do not know whether the two problems are fully equivalent, though this appears to be the case for a few small values of d and n. The combinatorial problem is plainly equivalent to the problem of the maximum number of intervals for n functions which are continuous but not necessarily of the form (2), and which have the property that any two of them are equal for at most $d-1$ values of x.

An obvious upper bound for $N_d(n)$ follows from the consideration that the pairs of integers

$$i(j),\ i(j+1), \quad \text{for } j = 1, 2, \cdots, N-1,$$

can include any given pair i_1, i_2 at most d times. Since the number of pairs i_1, i_2 with

$$1 \leqq i_1 \leqq n, \quad 1 \leqq i_2 \leqq n, \quad i_1 \neq i_2$$

is $n(n-1)$, it follows that

$$N_d(n) \leqq dn(n-1) + 1. \tag{6}$$

The problem of evaluating $N_d(n)$ is trivial when $d = 2$, for then there

is no subsequence a, b, a, and therefore any integer can occur only once. The longest sequences are simply the permutations of $1, 2, \cdots, n$, and we have

$$N_2(n) = n. \tag{7}$$

The case $d = 3$ is also simple. We prove:

THEOREM 1. *We have*

$$N_3(n) = 2n - 1, \tag{8}$$

and two examples of maximal sequences are

$$\begin{cases} 1, 2, 3, \cdots, n-1, n, n-1, \cdots, 3, 2, 1; \\ 1, 2, 1, 3, \cdots, 1, n-1, 1, n, 1. \end{cases} \tag{9}$$

For $d > 3$ the problem becomes much more difficult and appears to change its character. We shall concern ourselves mainly with the behaviour of $N_d(n)$ for fixed d and large n. As regards a lower bound for $N_d(n)$, we prove:

THEOREM 2. *We have*

$$N_d(n) \geqq (d^2 - 4d + 3)n - C(d) \quad \textit{if } d \textit{ is odd and } d > 3, \tag{10}$$

$$N_d(n) \geqq (d^2 - 5d + 8)n - C(d) \quad \textit{if } d \textit{ is even and } d > 4, \tag{11}$$

where $C(d)$ depends only on d. Also $N_4(n) \geqq 5n - C$.

As regards upper bounds, we prove:

THEOREM 3. *We have*

$$N_4(n) < 2n(1 + \log n), \tag{12}$$

and, for $d > 4$,

$$N_d(n) < An \exp\{B(\log n)^{\frac{1}{2}}\}, \tag{13}$$

where A, B depend only on d and

$$B = B(d) = 10(d \log d)^{\frac{1}{2}}. \tag{14}$$

2. *Proof of Theorem* 1. We give two proofs, based on different principles. Neither of them appears to be capable of extension to the case $d > 3$. In both proofs, $\mathcal{S}$ denotes a sequence of maximal length satisfying the conditions of the problem, that is, having no immediate repetition and containing no subsequence of the form a, b, a, b. We abbreviate $N_3(n)$ to $N(n)$.

First proof. We can suppose without loss of generality that the first term of $\mathcal{S}$ is 1. We can write $\mathcal{S}$ as

$$1, \mathcal{S}_1, 1, \mathcal{S}_2, \cdots, 1, \mathcal{S}_k, (1),$$

where each $\mathcal{S}_m$ is a sequence formed from the integers $2, 3, \cdots, n$, and where the final 1 may or may not occur. The sequences $\mathcal{S}_m$ are disjoint; for if an integer x occurred in two of them, there would be a subsequence 1, x, 1, x in $\mathcal{S}$. Thus if n_m denotes the number of distinct integers in $\mathcal{S}_m$, we have

$$n_1 + n_2 + \cdots + n_k \leqq n - 1.$$

Since $\mathcal{S}_m$ is a segment of $\mathcal{S}$, it satisfies the conditions of the problem, and therefore the number of terms in $\mathcal{S}_m$ is at most $N(n_m)$. It follows that

$$N(n) \leqq k + 1 + N(n_1) + \cdots + N(n_k).$$

By induction, starting from $N(1) = 1$, we obtain

$$N(n) \leqq k + 1 + (2n_1 - 1) + \cdots + (2n_k - 1) \leqq 2n - 1.$$

The fact that the particular sequences (9) have the desired property is obvious, and this proves (8).

Second proof. We begin with an observation, made to us by Mrs. Turan, that there is some one of the integers $1, 2, \cdots, n$ which occurs only once in $\mathcal{S}$. For if a is any integer which occurs twice in $\mathcal{S}$, so that

$$i(j_1) = a, \quad i(j_2) = a, \quad j_1 < j_2,$$

there must be some integer b which occurs between, say

$$i(j_3) = b, \quad j_1 < j_3 < j_2.$$

This integer b cannot occur as $i(j)$ for $j < j_1$ or $j > j_2$, for then we should have a subsequence b, a, b, a or a, b, a, b. If b occurs only once we have the result, and if not we can repeat the argument with b instead of a, and this process must terminate.

Now suppose, as we may without loss of generality, that n occurs only once in $\mathcal{S}$. If we delete the term n from $\mathcal{S}$, we obtain a sequence whose terms are formed from $1, 2, \cdots, n-1$ and which has no subsequence of the form a, b, a, b. This sequence may, however, have one immediate repetition, namely if the neighbours of n in $\mathcal{S}$ are equal:

$$\cdots, x, n', n, n', y, \cdots.$$

But this immediate repetition disappears if we delete also one of the two

terms n', since $x \neq n'$ and $y \neq n'$. Hence by deleting at most two terms from $\mathcal{S}$ we can obtain an admissible sequence whose terms are formed from $1, 2, \cdots, n-1$. It follows that

$$N(n) \leqq N(n-1) + 2,$$

and this again gives (8).

3. *Proof of Theorem 2. Suppose first that d is odd.* Let $\mathfrak{A}$ denote the sequence

$$1, 2, \cdots, n,$$

and let $\mathfrak{D}$ denote the sequence

$$n-1, n-2, \cdots, 2.$$

Then the sequence

$$\mathfrak{A}, \mathfrak{D}, \mathfrak{A}, \mathfrak{D}, \cdots, \mathfrak{A}, \mathfrak{D}, 1 \tag{15}$$

(which is symmetrical, in spite of its appearance) satisfies the conditions of the problem, provided each $\mathfrak{A}$ and $\mathfrak{D}$ is taken $(d-1)/2$ times. For if $a < b$, the successive pairs a, b in a subsequence $a, b, a, b, \cdots$ must have their a's in different $\mathfrak{A}$'s, assuming (as we may) that we take the last occurrence of each a before the corresponding b. Consequently there cannot be $(d+1)/2$ such pairs. By symmetry the same holds if $a > b$.

The sequence (15) has length $(d-1)(n-1)+1$. If $d > 3$ it can be expanded into a longer sequence, which is still admissible, as follows. We replace each element in $\mathfrak{A}$, say the first element 1, by

$$1, x, 1, x, \cdots, 1, x \quad \text{with } d-3 \text{ terms.}$$

Here x is an integer greater than n, and we use the same integer for all the elements of the first $\mathfrak{A}$ in (15). We do the same with each $\mathfrak{A}$ and $\mathfrak{D}$ in (15), but using a different new integer for each of them, and finally we replace the last term 1 in (15) by

$$1, t, 1, t, \cdots, 1, t \quad \text{to } d-3 \text{ terms.}$$

where t is the same new integer as that used to expand the last $\mathfrak{D}$. We now have a sequence with $n + (d-1)$ distinct terms, and of length

$$(d-3)\{(d-1)(n-1)+1\}.$$

We shall prove that this sequence satisfies the conditions of the problem, and it will follow that

$$N_d(n+d-1) \geqq (d-1)(d-3)(n-1) + (d-3),$$

which gives (10).

We have to prove that the expanded sequence contains no subsequence $a, b, a, b, \cdots$ with $d+1$ terms. No further proof is needed if $a \leqq n$ and $b \leqq n$, since then the subsequence is a subsequence of (15). The result is obviously true if $a > n$ and $b > n$, that is, if a and b both belong to the set $x, y, \cdots$ of additional integers, for then there is no subsequence of the form x, y, x. Thus we can suppose that either $a \leqq n$ and $b > n$ or $a > n$ and $b \leqq n$, and it will be enough to treat the former case. We replace b by y for ease of comparison with the construction.

In any subsequence

$$a, y, a, y, \cdots, a, y, \tag{16}$$

all the occurrences of y must be in the expansion of the same $\mathfrak{A}$ or $\mathfrak{D}$ in (15), or possibly in that of the final $\mathfrak{D}$ and 1. Except for the first y in (16), the a's which precede each y are in that same $\mathfrak{A}$ or $\mathfrak{D}$. The number of y's is therefore at most $\frac{1}{2}(d-3)+1$. Hence the length of the subsequence (16) is at most $d-1$, and this, when we allow for the possible occurrence of another a after (16), means a total length of at most d. Hence the expanded sequence has the desired property.

Suppose now that d is even. We start from the sequence

$$\mathfrak{A}, \mathfrak{D}, \mathfrak{A}, \mathfrak{D}, \cdots, \mathfrak{A}, \tag{17}$$

where $\mathfrak{A}$ occurs $\frac{1}{2}d$ times and $\mathfrak{D}$ occurs $\frac{1}{2}d-1$ times. The longest subsequence $a, b, a, b, \cdots$ in (17) has length d, or indeed only $d-1$ if $a > b$.

We expand (17) by replacing each element a in the first $\mathfrak{A}$ by

$$a, x, a, x, \cdots, a, x \quad \text{to } d-2 \text{ terms,}$$

where x is an integer greater than n. We apply the same treatment to the last $\mathfrak{A}$, using a different integer greater than n. We also expand the intermediate $\mathfrak{A}$'s and $\mathfrak{D}$'s, but here we replace each element a by

$$a, x, a, x, \cdots, x, \quad \text{to } d-4 \text{ terms,}$$

again using a different integer x for each $\mathfrak{A}$ and $\mathfrak{D}$. It can be proved, on the same lines as before, that the expanded sequence contains no subsequence $a, b, a, b, \cdots$ of $d+1$ terms.

The number of distinct terms in the expanded sequence is $n+d-1$, and the length is

$$> 2(d-2)n + (d-4)(d-3)(n-2) = (d^2-5d+8)n - 2(d-3)(d-4).$$

Hence

$$N_d(n+d-1) \geqq (d^2-5d+8)n - 2(d-3)(d-4),$$

and this gives (11). If $d=4$ we do not expand the intermediate $\mathfrak{C}$'s and $\mathfrak{D}$'s, and get $N_4(n) \geqq 5n - C$.

4. *Proof of* (12). Let $\mathcal{S}$ be a sequence of maximal length for $d=4$, this length being $N_4(n)$. Let $k(a)$ denote the number of times that a occurs in $\mathcal{S}$, for $a = 1, 2, \cdots, n$. Then

$$\sum_{a=1}^{n} k(a) = N_4(n). \tag{18}$$

If we delete a wherever it occurs in $\mathcal{S}$, we obtain a sequence formed from the $n-1$ integers other than a, and this sequence has no subsequence $a, b, a, b, \cdots$ of length greater than 4. But it may have immediate repetitions. To remove these, we must delete not only each occurrence of a but also one of the neighbours of a whenever these two neighbours are equal, as in the second proof of Theorem 1.

We now prove, for any a, that there are at most two occurrences of a, namely the first and the last, which can have equal neighbours. This is immediate, for in the contrary case we should have

$$\cdots, a, \cdots, x, a, x, \cdots, a, \cdots,$$

containing a subsequence a, x, a, x, a of 5 terms.

It follows that by deleting $k(a) + 2$ elements from $\mathcal{S}$ we can obtain an admissible sequence formed from $n-1$ distinct integers. Hence

$$N_4(n) \leqq N_4(n-1) + k(a) + 2.$$

Summing for $a = 1, \cdots, n$ and using (18), we obtain

$$nN_4(n) \leqq nN_4(n-1) + N_4(n) + 2n.$$

This can be written

$$\frac{N_4(n)}{n} - \frac{N_4(n-1)}{n-1} \leqq \frac{2}{n-1}.$$

Writing down a series of such equations and adding them, and noting that $N_4(2) = 4$, we obtain

$$\frac{N_4(n)}{n} - 2 \leqq 2\left(\tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{n-1}\right)$$
$$< 2\int_1^{n-1} t^{-1}dt = 2\log(n-1).$$

This proves (12).

5. We now prove (13), and begin with a simple lemma. Throughout

this section $\mathcal{S}$ will denote an admissible sequence for d and n, that is, a sequence formed from the integers $1, 2, \cdots, n$, with no immediate repetition, which contains no subsequence $a, b, a, b, \cdots$ of $d+1$ terms. These conditions imposed on $\mathcal{S}$ are unaffected by a permutation of $1, 2, \cdots, n$, and by choosing a suitable permutation, as in the lemma, we can simplify the later exposition.

LEMMA. *After a suitable permutation of* $1, 2, \cdots, n$, *the sequence* $\mathcal{S}$ *has the following properties*:

(i) *before any occurrence of any integer* m *in* $\mathcal{S}$ *there occur all integers less than* m;

(ii) $\mathcal{S}$ *contains no subsequence*

$$m, a, b, a, b, \cdots, \text{ to } d \text{ terms altogether,} \tag{19}$$

with

$$m \geqq b > a. \tag{20}$$

Proof. We take the first term of $\mathcal{S}$ to be 1, the second term to be 2, the next term other than 1 and 2 to be 3, and so on, numbering the terms in the order of their first appearance in $\mathcal{S}$. Plainly (i) holds.

Suppose $\mathcal{S}$ has a subsequence of the form (19), where m, a, b satisfy (20). Then before the first term, m, in (19), or possibly coinciding with it, there occurs a term b, since $b \leqq m$. Before this term b there occurs a term a, since $a < b$. But then there is a subsequence $a, b, a, b, \cdots$ to $d+1$ terms, contrary to hypothesis. This proves the lemma.

We remark that (ii) implies the original hypothesis that $\mathcal{S}$ contains no subsequence $a, b, a, b, \cdots$ to $d+1$ terms, since such a sequence always contains a sequence of d terms with the first term greater than the second, and this is excluded by (19) with $m = b$.

Proof of (13). For any integer m with $1 < m < n$ we pick out the first occurrence of m in $\mathcal{S}$ and dissect $\mathcal{S}$ into

$$\mathcal{S}', m, \mathcal{S}'', \tag{21}$$

so that every term in $\mathcal{S}'$ is one of $1, 2, \cdots, m-1$.

We write $\mathcal{S}''$ as

$$\mathcal{S}_1^{(1)}, a_1, \mathcal{S}_1^{(2)}, a_1, \cdots, \mathcal{S}_1^{(r_1)}, a_1, \mathcal{S}_2^{(1)}, a_2, \cdots, a_k, \mathcal{S}_k^{(r_k)}, a_k, \mathcal{T}, \tag{22}$$

where $a_1, \cdots, a_k$ are all the terms not exceeding m that occur in $\mathcal{S}''$, and all the terms of the sequences $\mathcal{S}_i^{(j)}$ and $\mathcal{T}$ are integers greater than m. Note

that the integers $a_1, \cdots, a_k$ are not necessarily distinct, though $a_i \neq a_{i+1}$ as a consequence of our choice of notation.

The sequence $\mathcal{S}'$ consists of terms each of which is one of $1, 2, \cdots, m-1$, and is an admissible sequence. Hence

$$(23) \qquad L(\mathcal{S}') \leqq N_d(m-1),$$

where $L(\mathcal{S}')$ denotes the length of $\mathcal{S}'$.

The sequence $a_1, a_2, \cdots, a_k$ has each term less than or equal to m and has no immediate repetition. It also contains no subsequence $a, b, a, b, \cdots$ of d terms, for this would necessarily contain a similar sequence of $d-1$ terms with the first term less than the second, and this, preceded by m, would contradict (ii) of the lemma. Hence

$$(24) \qquad k \leqq N_{d-1}(m).$$

The sequence

$$\mathcal{S}_1^{(1)}, \mathcal{S}_1^{(2)}, \cdots, \mathcal{S}_1^{(r_1)}, \mathcal{S}_2^{(1)}, \cdots, \mathcal{S}_2^{(r_2)}, \cdots, \mathcal{S}_k^{(1)}, \cdots, \mathcal{S}_k^{(r_k)}, \mathcal{T}$$

has all its terms greater than m, and is an admissible sequence except for possible immediate repetitions. These occur only when the last term of one of the above sequences is the same as the first term of the next. They can be removed by deleting at most $\sum r_i$ terms at the ends of the sequences. Hence

$$(25) \qquad \sum_{i=1}^{k} \sum_{j=1}^{r_i} L(\mathcal{S}_i^{(j)}) + L(\mathcal{T}) \leqq N_d(n-m) + \sum_{i=1}^{k} r_i.$$

It remains to estimate $\sum r_i$. For this we consider only the sequences $\mathcal{S}_i^{(j)}$ with $j > 1$. None of them can be empty, since otherwise there would be an immediate repetition of some a_i in (22). We select from each of these sequences a term $x_i^{(j)}$. Among the terms

$$(26) \qquad x_i^{(2)}, x_i^{(3)}, \cdots, x_i^{(r_i)},$$

for given i, the same integer cannot occur more than $\frac{1}{2}d$ times, since otherwise there would be a subsequence

$$x, a_i, x, a_i, \cdots, x, a_i$$

of more than d terms. It follows that the number of distinct integers among (26) is at least $2(r_i-1)/d$.

Let $\mathfrak{X}_i$ be a subsequence of (26) containing s_i distinct terms, where $s_i \geqq 2(r_i-1)/d$. The sequence

$$(27) \qquad \mathfrak{X}_1, \mathfrak{X}_2, \cdots, \mathfrak{X}_k$$

is admissible for d, except for possible immediate repetitions. Since the terms of each $\mathfrak{X}_i$ are distinct among themselves, all immediate repetitions can be removed by deleting at most $k-1$ terms. Since all the terms in (27) are greater than m, and the total number of terms is $\sum s_i$, we have

$$\sum_{i=1}^{k} s_i \leqq N_d(n-m)+k-1.$$

It follows that

$$2d^{-1}\sum_{i=1}^{k}(r_i-1)\leqq N_d(n-m)+k-1,$$

whence

$$\sum_{i=1}^{k} r_i<\tfrac{1}{2}dN_d(n-m)+(\tfrac{1}{2}d+1)k. \tag{28}$$

By (23), (24), (25), (28) we have

$$\begin{aligned} L(\mathcal{S}) &\leqq L(\mathcal{S}')+1+\sum_{i=1}^{k} r_i+\sum_{i=1}^{k}\sum_{j=1}^{r_i} L(\mathcal{S}_i^{(j)})+L(\mathcal{T}) \\ &\leqq N_d(m-1)+1+\sum_{i=1}^{k} r_i+N_d(n-m)+\sum_{i=1}^{k} r_i \\ &\leqq N_d(m-1)+(d+1)N_d(n-m)+(d+2)k \\ &\leqq N_d(m-1)+(d+1)N_d(n-m)+(d+2)N_{d-1}(m). \end{aligned}$$

Taking $\mathcal{S}$ to be a maximal sequence, we obtain the inductive inequality

$$N_d(n)\leqq N_d(m)+(d+1)N_d(n-m)+(d+2)N_{d-1}(m). \tag{29}$$

Suppose that $d\geqq 5$ and that

$$N_{d-1}(m)<A_1 m\exp(B_1\sqrt{\log m}) \tag{30}$$

for all m, where

$$B_1=10\{(d-1)\log(d-1)\}^{\frac{1}{2}} \tag{31}$$

and A_1 depends only on d. This is a legitimate assumption when $d=4$, by (12).

Choose A sufficiently large to ensure that the inequality

$$N_d(m)<Am\exp(B\sqrt{\log m}), \tag{32}$$

where

$$B=10\{d\log d\}^{\frac{1}{2}}, \tag{33}$$

holds for all $m\leqq n_0$, where $n_0=n_0(d)$ will be chosen later (in a manner which does not depend on the choice of A). Suppose also that

$$A>2(d+2)A_1. \tag{34}$$

Now suppose that $n > n_0$ and that (32) holds for all $m < n$; we have to prove that it holds for $m = n$. Define $C = B - B_1$. Let h be the integer defined by

$$h - 1 < n \exp(-C\sqrt{\log n}) \leqq h. \tag{35}$$

We suppose n_0 chosen sufficiently large to ensure that $1 < h < n$. By (29),

$$\begin{aligned} N_d(n) &\leqq N_d(n-h) + (d+1)N_d(h) + (d+2)N_{d-1}(n-h) \\ &< A(n-h)\exp(B\sqrt{\log n}) + (d+1)Ah\exp(B\sqrt{\log h}) \\ &\qquad + (d+2)A_1(n-h)\exp(B_1\sqrt{\log n}). \end{aligned}$$

This will be less than $An\exp(B\sqrt{\log n})$, thus giving the desired conclusion, provided that

$$\begin{aligned} Ah\exp(B\sqrt{\log n}) &> (d+1)Ah\exp(B\sqrt{\log h}) \\ &\qquad + (d+2)A_1 n\exp(B_1\sqrt{\log n}). \end{aligned}$$

Since $n/h \leqq \exp((B - B_1)\sqrt{\log n})$ by (35), it will suffice if

$$A > (d+1)A\exp(-B\sqrt{\log n} + B\sqrt{\log h}) + (d+2)A_1.$$

By (34), this will hold if

$$1 > 2(d+1)\exp(-B\sqrt{\log n} + B\sqrt{\log h}).$$

Now, by (35),

$$\sqrt{\log n} - \sqrt{\log h} > \sqrt{\log n} - \sqrt{(\log 2n - C\sqrt{\log n})} > \tfrac{1}{3}C,$$

provided n_0 is sufficiently large. Hence it suffices if

$$BC > 3\log 2(d+1).$$

By (31), (33),

$$\begin{aligned} C = B - B_1 &= 10\{(d\log d)^{\frac{1}{2}} - ((d-1)\log(d-1))^{\frac{1}{2}}\} \\ &> 5d^{-\frac{1}{2}}(\log d)^{\frac{1}{2}}. \end{aligned}$$

Hence

$$BC > 50\log d,$$

and this amply suffices. This completes the proof of (13).

Note added in proof. Since this paper was written we have improved on the results of Theorems 2 and 3.

TRINITY COLLEGE, CAMBRIDGE, ENGLAND
AND
OHIO STATE UNIVERSITY, COLUMBUS, OHIO.

MEAN CONVERGENCE OF EXPANSIONS IN LAGUERRE AND HERMITE SERIES.

By Richard Askey[1] and Stephen Wainger.[2]

1. Introduction. A number of authors have studied the problem of mean convergence in L^p for expansions of functions in orthogonal functions. Of particular interest are the results of Pollard and Wing, [6], [7], [9], for orthogonal polynomials. Wing has shown the following.

Let $\{R_n(x)\}$ be polynomials orthonormal with respect to a weight function $w(x)$, $w(x)$ defined on a bounded interval. Then under various technical conditions on w the following mean convergence theorem holds. Let $f \in L^p$, $4/3 < p < 4$. Then if

$$f(x) \sim \sum a_n w^{\frac{1}{2}}(x) R_n(x),$$

i. e.

$$a_n = \int f(x) w^{\frac{1}{2}}(x) R_n(x)\, dx$$

and if

$$S_n(x) = \sum_{k=0}^{n} a_k w^{\frac{1}{2}}(x) R_n(x)$$

then

$$\| S_n - f \|_p \to 0.$$

The most important class of w in Wing's result is $w(x) = (1-x)^\alpha (1+x)^\beta$ and then the corresponding polynomials are the Jacobi polynomials. We obtain the analogue of this result for the other classical polynomials, $L_n^\alpha(x)$, the Laguerre polynomials, and $H_n(x)$, the Hermite polynomials. There are added problems because the intervals for which these polynomials are orthogonal are no longer bounded.

Let $L_n^\alpha(x)$, $\alpha \geqq 0$, be defined by

$$\sum L_n^\alpha(x) r^n = (1-r)^{-\alpha-1} \exp\left(-\frac{xr}{1-r}\right).$$

Then the functions

$$\mathscr{L}_n^\alpha(x) = x^{\alpha/2} r_n \exp(-x/2) L_n^\alpha(x)$$

Received August 6, 1964.
Revised January 8, 1965.
[1] Supported in part by N. S. F. grant GP-3483.
[2] Supported in part by N. S. F. grant GP-1645.

are orthonormal on $(0,\infty)$. Here $r_n = \{\Gamma(n+\alpha+1)/n!\}^{-\frac{1}{2}}$. Let $L^p(0,\infty)$ be the space of measurable functions such that $\| f \|_p = [\int_0^\infty | f(x) |^p \, dx]^{1/p}$ is finite. The main result of this paper is the following theorem.

THEOREM 1. *Let $f(x)$ be in $L^p(0,\infty)$, $4/3 < p < 4$. Define*

$$a_n = \int_0^\infty \mathcal{L}_n{}^\alpha(x) f(x)\, dx$$

and set

$$S_n = \sum_{k=0}^{n} a_k \mathcal{L}_k{}^\alpha(x).$$

Then $\| S_n - f \|_p \to 0$ as $n \to \infty$.

As usual results of this type are proved by first proving uniform L^p boundedness of the partial sums. That is, the essential part of the proof is contained in Theorem 1′.

THEOREM 1′. *Let f, $\mathcal{L}_n{}^\alpha$, a_n, s_n, and p be as in Theorem 1. Then there exists an A_p such that*

$$\| S_n \|_p \leqq A_p \| f \|_p, \qquad (n = 0, 1, \cdots),$$

where A_p is independent of $f \in L^p$.

We also consider the corresponding problem for Hermite series. We define the Hermite polynomial $H_n(x)$ by

$$\sum H_n(x) r^n = \exp(2xr - r^2).$$

The functions $\mathcal{H}_n(x) = \exp(-x^2/2)\pi^{-1/4} 2^{-n/2}[n!]^{-1/2} H_n(x)$ are orthonormal on $(-\infty,\infty)$. $L^p(-\infty,\infty)$ is defined similary to the above. The following theorems are analogous to Theorems 1 and 1′ above.

THEOREM 2. *Let f be in $L^p(-\infty,\infty)$, $4/3 < p < 4$. Define*

$$a_n = \int_0^\infty f(x) \mathcal{H}_n(x)\, dx$$

and set

$$S_n = \sum_{k=0}^{n} a_k \mathcal{H}_k(x).$$

Then $\| S_n - f \|_p$ tends to 0 as n approaches infinity.

THEOREM 2′. *Let f, $\mathcal{H}_n$, a_n, s_n, and p be as in Theorem 2. Then there is an A_p such that*

$$\| S_n \|_p \leqq A_p \| f \|_p, \qquad (n = 0, 1, \cdots),$$

where A_p is independent of $f \in L^p$.

Remark 1. We shall show that the above theorems are false for $1 \leqq p \leqq 4/3$ and $p \geqq 4$.

Remark 2. The following questions might seem more natural than Theorems 1 and 2. Let $|||f|||_p = (\int_0^\infty x^\alpha e^{-x} |f(x)|^p\,dx)^{1/p}$. If $|||f|||_p < \infty$, define $r_n^2 c_n = \int_0^\infty x^\alpha e^{-x} f(x) L_n^\alpha(x)\,dx$. Set $S_n = \sum_{k=0}^n c_k L_k^\alpha(x)$. Then does $|||S_n - f|||_p \to 0$ as $n \to \infty$? Pollard [7] has shown that the answer to this question and the corresponding question for Hermite polynomials is "no" unless $p = 2$.

We wish to thank the referee for comments which helped in the exposition of this work.

2. Preliminary material. The proof of Theorems 1′ and 2′ depends upon combining an ingenious device of Pollard [6] with recent asymptotic estimates for Laguerre and Hermite polynomials by Erdélyi and Skovgaard. See [2] and [3]. Pollard showed that in studying mean convergence problems one can consider $|R_n(x) - R_{n-2}(x)|$ rather than $|R_n(x) - R_{n-1}(x)|$. (The later is what one is led to directly from the Christoffel Darboux formula.) It is well known that $R_n - R_{n-2}$ behaves better than $R_n - R_{n-1}$ if the R_n are Jacobi polynomials. We shall show that the recent asymptotic formulas provide sufficiently good information for $\mathcal{L}_n^\alpha - \mathcal{L}_{n-2}^\alpha$ and $\mathcal{H}_n - \mathcal{H}_{n-2}$.

For the convenience of the reader we state the theorems of Erdélyi and Skovgaard, [2] and [3].

THEOREM A. (Erdélyi)

(i) *Let* $0 \leqq x \leqq b\nu$, $n \geqq n_0$, $b < 1$.

$$L_n^\alpha(x) = (n!)^{-1}\Gamma(n+\alpha+1)2^{(\alpha-1/2)}\nu^{(1-\alpha)/2}x^{-(\alpha+1)/2}$$
$$\cdot e^{(x/2)}\left(\frac{\psi}{\psi'}\right)^{\frac12}\left\{J_\alpha(\nu\psi) + O\left[\frac{1}{\nu}\left(\frac{x}{\nu - x}\right)^{\frac12}\cdot \tilde{J}_\alpha(\nu\psi)\right]\right\}.$$

(ii) *Let* $0 < a\nu \leqq x$, $n \geqq n_0$, $a > 0$.

$$L_n^\alpha(x) = \frac{(-1)^n}{n!}\left(\frac{\pi}{-\phi'}\right)^{1/2}2^{5/6}N^{N+1/6}e^{-N}$$
$$\cdot x^{-(\alpha+1)/2}\cdot e^{(x/2)}\left\{Ai(-\nu^{2/3}\phi) + O\left[\frac{1}{x}\widetilde{Ai}(-\nu^{2/3}\phi)\right]\right\}.$$

In (i) and (ii) above $N = n + (\alpha+1)/2$ and $\nu = 4N = 4n + 2\alpha + 2$. $\psi = \psi(t)$ satisfies

$$[\psi'(t)]^2 = \tfrac14\left(\frac{1}{t} - 1\right).$$

$\phi = \phi(t)$ is the unique real continuous solution of

$$\phi(t)[\phi'(t)]^2 = \tfrac{1}{4}(\frac{1}{t} - 1), \quad 0 < t < \infty.$$

$$t = x/\nu.$$

$$\psi(t) = \tfrac{1}{2}[(t - t^2)^{\frac{1}{2}} + \sin^{-1} t^{\frac{1}{2}}], \quad 0 \leqq t < 1.$$

$$\phi(t) = \begin{cases} [3\beta(t)/2]^{2/3}, & 0 < t \leqq 1. \\ -[3\gamma(t)/2]^{2/3}, & 1 < t. \end{cases}$$

$$\beta(t) = \tfrac{1}{2}[\cos^{-1} t^{\frac{1}{2}} - (t - t^2)^{\frac{1}{2}}].$$

$$\gamma(t) = \tfrac{1}{2}[(t^2 - t)^{\frac{1}{2}} - \cosh^{-1} t^{\frac{1}{2}}].$$

$$\widetilde{J}_\alpha(u) = \begin{cases} J_\alpha(u), \ u \text{ sufficiently small.} \\ (|J_\alpha(u)|^2 + |Y_\alpha(u)|^2)^{\frac{1}{2}} \text{ otherwise.} \end{cases}$$

$$\widetilde{Ai}(z) = \begin{cases} Ai(z) \text{ if } z \geqq 0. \\ (|Ai(z)|^2 + |Bi(z)|^2)^{\frac{1}{2}} \text{ if } z \leqq 0. \end{cases}$$

Y_α and J_α are Bessel functions of order α. See [1]. Ai and Bi are Airy integrals. See [4].

We shall need certain information about ϕ and ψ. Erdélyi [2] points out that $\psi(t) = t^{\frac{1}{2}}[1 + O(t)]$, and $\psi'(t)/\psi(t) = [2t]^{-1}[1 + O(t)]$.

We have to analyse $\phi(t)$ ourselves.

$$\beta'(t) = -(\tfrac{1}{2})(1 - t)^{\frac{1}{2}}(t)^{-\frac{1}{2}}.$$

So

$$\beta'(t) = -(\tfrac{1}{2})(1 - t)^{\frac{1}{2}}(1 + \sum_{n=1}^{\infty} b_n(1 - t)^n)$$

with $b_n \geqq 0$ and with the infinite series converging uniformly for $\delta \leqq t \leqq 1$. Therefore $\beta(t) = \frac{1}{3}(1 - t)^{3/2}(1 + \sum_{n=1}^{\infty} c_n(1 - t)^n)$ with $c_n \geqq 0$. Thus

$$\phi(t) = (\tfrac{1}{2})^{2/3}(1 - t)\sum_{n=0}^{\infty} d_n(1 - t)^n.$$

$\sum_{n=0}^{\infty} d_n(1 - t)^n$ converges to a positive function for $\delta \leqq t \leqq 1$. In a similar way one can show

$$\phi(t) = (\tfrac{1}{2})^{2/3}(t - 1)\sum_{n=0}^{\infty} g_n(t - 1)^n$$

where $\sum_{n=0}^{\infty} g_n(t - 1)^n$ converges to a positive function for $1 \leqq t \leqq 1 + \lambda$ for

some $\lambda > 0$. Finally $At \leqq \gamma(t) \leqq t/2$ for some constant A and $t \geqq (1+\lambda)$. So in this region $(At)^{2/3} \leqq -\phi(t) \leqq (t/2)^{2/3}$.

To use Theorem A we need certain information about Bessel functions and Airy integrals.

$|J_\alpha(z)| = O(|z|^\alpha)$ as $z \to 0$. ([1], p. 4, equation 2).

$J'_\alpha(z) = \frac{1}{2}[J_{\alpha-1}(z) - J_{\alpha+1}(z)]$. ([1], p. 12, equation 57).

$|J_\alpha(z)| = O(|z|^{-\frac{1}{2}})$ as $|z| \to \infty$, z real. ([1], p. 85, equation 3).

$|Y_\alpha(z)| = O(|z|^{-\frac{1}{2}})$ z real, $|z| \to \infty$. ([1], p. 85, equation 4).

$Ai(z)$ and $Bi(z)$ are entire functions. For $u > 0$.

$$|Ai(-u)| = O(u^{-\frac{1}{4}}),$$

and

$$|Bi(-u)| = O(u^{-\frac{1}{4}}).$$

$$|Ai(u)| = O(u^{-\frac{1}{4}} \exp(-2/3u^{3/2})),$$

see [4], pp. 508-511. We must also use the fact that

$$|Ai'(z)| = O[(1 + |z|^{\frac{1}{2}})|\widetilde{Ai}(z)|].$$

See [2].

Our corollaries are somewhat tedious but rather straightforward consequences of Theorem A and the mean value theorem. We list the results in tabular form. The results hold for $n \geqq n_0$.

Interval of x	$\mathscr{L}_n^\alpha(x)$	$\mathscr{L}_n^\alpha(x) - \mathscr{L}_{n+2}^\alpha(x)$
$0 \leqq x \leqq 1/n$	$O(x^{\alpha/2} n^{\alpha/2})$	$O(x^{\alpha/2} n^{\alpha/2-1})$
$1/n \leqq x \leqq \delta n$	$O(n^{-\frac{1}{4}} x^{-\frac{1}{4}})$	$O(n^{-\frac{3}{4}} x^{\frac{1}{4}})$
$\delta n \leqq x \leqq \nu - \nu^{\frac{1}{3}}$	$O(n^{-\frac{1}{4}}(\nu - x)^{-\frac{1}{4}})$	$O(n^{-\frac{3}{4}}(\nu - x)^{\frac{1}{4}})$
$\nu - \nu^{\frac{1}{3}} \leqq x \leqq \nu + \nu^{\frac{1}{3}}$	$O(n^{-\frac{1}{3}})$	$O(n^{-2/3})$
$\nu + \nu^{\frac{1}{3}} \leqq x \leqq (1+\lambda)\nu$	$O\{[(n(x-\nu))^{-\frac{1}{4}} \cdot \exp(-\eta(x-\nu)^{\frac{3}{2}}\nu^{-\frac{1}{2}})\}$	$O\{n^{-\frac{3}{4}}(x-\nu)^{\frac{1}{4}} \cdot \exp[-\eta(x-\nu)^{\frac{3}{2}}\nu^{-\frac{1}{2}}]\}$
$(1+\lambda)\nu \leqq x$	$O(\exp(-\xi x))$	$O(\exp(-\xi x))$

In the above table δ and λ are sufficiently small positive numbers. $0 < \eta < 2/3$, and $0 < \xi < 1/2$.

One consequence of the above estimates is that for $4/3 < p < 4$

$$\| \mathscr{L}_n^\alpha(x) \|_p = O(n^{-(1/2-1/p)})$$

so

$$\| \mathcal{L}_n{}^{\alpha}(x) \|_p \| \mathcal{L}_n{}^{\alpha}(x) \|_q = O(1)$$

for $4/3 < p < 4$ and $1/p + 1/q = 1$. We shall need this fact later.

We now state the analogue of Theorem A for Hermite polynomials.

THEOREM B. (Skovgaard)

$$\mathcal{H}_n(N^{\frac12}x) = (2\pi)^{\frac14} N^{n/2+1/6} \exp[\tfrac14 N(x^2 - \tfrac12)]$$

$$|\phi'|^{-1/2}\{Ai(-N^{2/3}\phi) + O\,\frac{\widetilde{Ai}(-N^{2/3}\phi)}{n(1+x^2)}\},$$

valid uniformly for

$$-1 < a \leqq x < \infty$$

as $n \to \infty$. *Here* $N = 2n+1$.

$$\tfrac23 \phi^{\frac32}(x) = \int_x^1 (1-t^2)^{\frac12}\,dt; \qquad a \leqq x \leqq 1.$$

$$\tfrac23 (-\phi)^{\frac32} = \int_1^x (t^2-1)^{\frac12}\,dt; \qquad 1 \leqq x < \infty.$$

From Theorem B, we derive the following table.

Interval of x	$\mathcal{H}_n(x)$	$\mathcal{H}_n(x) - \mathcal{H}_{n+2}(x)$
$aN^{1/2} \leqq x \leqq N^{1/2} - N^{-1/6}$	$O\{N^{-1/8}(N^{1/2}-x)^{-1/4}\}$	$O\{N^{-3/8}(N^{1/2}-x)^{1/4}\}$
$N^{1/2} - N^{-1/6} \leqq x \leqq N^{1/2} + N^{-1/6}$	$O\{N^{-1/12}\}$	$O\{N^{-5/12}\}$
$N^{1/2} + N^{-1/6} \leqq x \leqq (2N)^{1/2}$	$O\{N^{-1/8}(x-N^{1/2})^{-1/4}$ $\exp[-\xi N^{1/4}(x-N^{1/2})^{3/2}]\}$	$O\{N^{-3/8}(x-N^{1/2})^{1/4}$ $\exp[-\xi N^{1/4}(x-N^{1/2})$
$(2N)^{1/2} \leqq x$	$O\{\exp(-\xi x^2)\}$	$O\{\exp(-\xi x^2)\}$

In the above table ξ is some positive number. The table may be extended to negative values of x by the formula $H_n(-x) = (-1)^n H_n(x)$. [1], p. 193, equation 14.

Finally we remark that our table implies

$$\| \mathcal{H}_n(x) \|_p \cdot \| \mathcal{H}_n(x) \|_q$$

is bounded for $4/3 < p < 4$ and $1/p + 1/q = 1$.

Important tools for us will be Hardy's inequality and the theory of the

Hilbert transform. For Hardy's inequality see p. 20 of [10], and for the theory of the Hilbert transform see [5]. We summarize relevant facts below.

The simplest form of Hardy's inequality is the following:

Let $f(x)$ be in $L^p[0,\infty]$ with $1 < p < \infty$. Define

$$F(x) = \frac{1}{x}\int_0^x f(t)\,dt.$$

Then $F(x)$ is in L^p, and

$$\| F \|_p \leqq A_p \| f \|_p.$$

Hardy's inequality remains true if certain weight factors are added. In particular if $4/3 < p < 4$, and if

$$\int_0^\infty | f(x) |^p\, dx < \infty,$$

then

$$F_0(x) = \frac{1}{x^{\pm\frac{1}{4}}x} \int_0^x f(t)\, t^{\pm\frac{1}{4}}\, dt$$

is in L^p, and

$$\| F_0(x) \|_p \leqq A_p \| f \|_p.$$

(The $+$ or $-$ must be taken the same in both occurrences.)

By changing variables, we see that the same conclusion holds if we replace $F_0(x)$ by $F_a(x)$ where

$$F_a(x) = \frac{1}{| x - a |^{\pm\frac{1}{4}}x} \int_0^\infty f(t)\, | t - a |^{\pm\frac{1}{4}}\, dt,$$

where a is any real number.

Either by using similar methods or by making a change of variables $x \to 1/x$, one can see the same conclusions hold if we redefine $F(x)$ by

$$F(x) = \int_x^\infty f(t)\, t^{-1}\, dt$$

and redefine

$$F_a(x) = | x - a |^{\pm\frac{1}{4}} \int_x^\infty f(t)\, t^{-1} t^{\mp\frac{1}{4}}\, dt.$$

We now discuss the Hilbert transform. Let $f(x)$ be in $L^p[0,\infty)$, $1 < p < \infty$. Then

$$\tilde{f}(x) = \lim_{\epsilon \to 0+} \int_{|x-y|>\epsilon} \frac{1}{x - y}\, f(y)\, dy$$

exists almost everywhere and is in L^p. Moreover

$$\| \tilde{f}(x) \|_p \leqq A_p \| f(x) \|_p.$$

As in the case of Hardy's inequality certain weighted results are true. In particular if $4/3 < p < 4$, and if f is in L^p,

$$\tilde{f}_a(x) = \lim_{\epsilon \to 0+} | x - a |^{\pm\frac{1}{4}} \int_{|x-y|>\epsilon} \frac{| y - a |^{\mp\frac{1}{4}}}{x - y} f(y)\, dy$$

exists almost everywhere and is in L^p. Moreover

$$\| \tilde{f}_a(x) \|_p \leqq A_p \| f(x) \|_p.$$

Finally, if one combines the theory of the Hilbert transform with Hardy's inequality, one sees that the range of integration in the definition of $\tilde{f}$ and $\tilde{f}_a$ may be changed from $| x - y | > \epsilon$ to $| x - y | > \epsilon$ and $\omega x < y < \Omega x$ for fixed ω and Ω with $0 < \omega < 1 < \Omega$. The A_p will then also depend on ω and Ω.

3. **Proofs.** We now prove Theorem 1′. Let $f(x)$ be in $L^p[0, \infty]$, and $n \geqq n_0$ (i. e., the n_0 of Theorem A.).

$$S_n(x) = \sum_{k=0}^{n} \mathcal{L}_n{}^\alpha(x) \int_0^\infty f(y) \mathcal{L}_n{}^\alpha(y)\, dy$$
$$= \int_0^\infty f(y) D_n(x, y)\, dy.$$

$D_n(x, y) = \sum_{k=0}^{n} \mathcal{L}_n{}^\alpha(x) \ \mathcal{L}_n{}^\alpha(y)$. Now

$$1) \qquad D_n(x, y) = \frac{(n+1)!}{\Gamma(n+\alpha+1)} \frac{(xy)^{\alpha/2}}{x-y} \cdot e^{-(x/2)} e^{-(y/2)}$$
$$\cdot \{ \mathcal{L}_n{}^\alpha(x) \ \mathcal{L}_{n+1}{}^\alpha(y) - \mathcal{L}_{n+1}{}^\alpha(x) \ \mathcal{L}_n{}^\alpha(y) \}.$$

[1], p. 188, equation 9.

We now use Pollard's device of considering

$$\frac{D_n + D_{n+1}}{2} = D_n + \tfrac{1}{2} (\mathcal{L}_{n+1}{}^\alpha(x) \ \mathcal{L}_{n+1}{}^\alpha(y)).$$

Adding together the closed form expressions 1) for D_n and D_{n+1}, we find

$$D_n(x, y) = \frac{n}{x-y} \{ \mathcal{L}_{n+1}{}^\alpha(x) \ [\mathcal{L}_{n+2}{}^\alpha(y) - \mathcal{L}_n{}^\alpha(y)]$$
$$+ \mathcal{L}_{n+1}{}^\alpha(y) \ [\mathcal{L}_n{}^\alpha(x) - \mathcal{L}_{n+2}{}^\alpha(x)] \}$$
$$+ \frac{b_n}{x-y} [\mathcal{L}_{n+1}{}^\alpha(x) \ \mathcal{L}_{n+2}{}^\alpha(y) + \mathcal{L}_n{}^\alpha(x) \ \mathcal{L}_{n+1}{}^\alpha(y)$$
$$- \mathcal{L}_n{}^\alpha(x) \ \mathcal{L}_{n+1}{}^\alpha(y) - \mathcal{L}_{n+2}{}^\alpha(x) \mathcal{L}_{n+1}{}^\alpha(y)] - \tfrac{1}{2} \mathcal{L}_{n+2}{}^\alpha(x) \ \mathcal{L}_{n+2}{}^\alpha(y).$$

$$| b_n | = O(1).$$

We therefore may write

$$S_n = M_n + M_n' + E_{n,1} + E_{n,2}.$$

$$M_n = n\mathcal{L}_{n+1}^{\alpha}(x) \int_0^{\infty} \frac{f(y)}{x-y}[\mathcal{L}_n^{\alpha}(y) - \mathcal{L}_{n+2}^{\alpha}(y)]\,dy.$$

$$M_n' = n[\mathcal{L}_n^{\alpha}(x) - \mathcal{L}_{n+2}^{\alpha}(x)] \int_0^{\infty} \frac{f(y)}{x-y}\mathcal{L}_{n+1}^{\alpha}(y)\,dy.$$

$$E_{n,1} = b_n \int_0^{\infty} \frac{f(y)}{x-y}\{\mathcal{L}_{n+1}^{\alpha}(x)\,\mathcal{L}_{n+2}^{\alpha}(y) + \mathcal{L}_n^{\alpha}(x)\mathcal{L}_{n+1}^{\alpha}(y) - \mathcal{L}_n^{\alpha}(x)\mathcal{L}_{n+1}^{\alpha}(y) - \mathcal{L}_{n+2}^{\alpha}(x)\,\mathcal{L}_{n+1}^{\alpha}(y)\}\,dy.$$

$$E_{n,2} = \tfrac{1}{2} \int_0^{\infty} f(y)\mathcal{L}_n^{\alpha}(y)\,\mathcal{L}_n^{\alpha}(x)\,dy.$$

The singular integrals which appear above should be interpreted in the Cauchy principal value sense.

It follows from the estimates of Section 1 that the $\mathcal{L}_n^{\alpha}(x)$ are uniformly bounded. Therefore

$$\| E_{n,1} \|_p \leqq A_p \| f \|_p.$$

See [5]. Also

$$\| E_{n,2} \|_p \leqq A_p \| f \|_p. \quad (4/3 < p < 4)$$

This follows by Holder's inequality and the fact that

$$\| \mathcal{L}_n^{\alpha}(x) \|_p \, \| \mathcal{L}_n^{\alpha}(x) \|_q = O(1)$$

$$1/p + 1/q = 1, \quad 4/3 < p < 4.$$

(This was discussed in Section 2.)

We must make one more simple remark. Namely an easy duality argument shows that it suffices to consider either M_n and M_n'. We examine M_n. We divide the y-range of integration and the x-range each into 6 parts. These parts will be

$$0 \leqq x \leqq 1/n, \quad 1/n \leqq x \leqq \delta n, \quad \delta n \leqq x \leqq n - n^{\frac{1}{3}},$$
$$\nu - n^{\frac{1}{3}} \leqq x \leqq \nu + n^{\frac{1}{3}}, \quad \nu + n^{\frac{1}{3}} \leqq x \leqq (1+\lambda)n, \text{ and } (1+\lambda)n \leqq x.$$

ν, δ, and λ are as in the table of estimates on Laguerre polynomials. We thus express M_n as a sum of 36 integrals. We use the tabular estimates to study these integrals. Some of these terms can be seen to be bounded in L^p by a direct calculation, i.e., for $4/3 < p < 4$. Other integrals can be expressed in the form

2) $$\phi(x,n) \int \frac{f(y)}{x-y}\, \psi(y,n)\, dy$$

or

3) $$|x|^{-(\frac{1}{4})}\phi(x,n) \int \frac{f(y)}{x-y}\, \psi(y,n)\, |y|^{\frac{1}{4}}\, dy$$

or

3A) $$|x-\nu|^{-(\frac{1}{4})}\phi(x,n) \int \frac{f(y)}{y-x}\, \psi(y,n)\, |y-\nu|^{\frac{1}{4}}\, dy$$

or

finally the range of integration might have to be divided into two parts, one part of which can be handled directly and another part of which is of the form 2, 3, or 3A. In these formulas ϕ and ψ are bounded functions differing for different ranges of integration. The only dependence of the ranges of integration on x is of the form $\omega x < y < \Omega x$ with $0 \leqq \omega < 1 < \Omega \leqq \infty$. Integrals of the form 2), 3), and 3A) are considered in § 2 and all have L^p norm less than or equal to $A_p \| f \|_p$, $4/3 < p < 4$.

Of the above mentioned thirty-six integrals we shall discuss three typical ones.

Consider first the term $\delta n \leqq y \leqq \nu - \nu^{\frac{1}{3}}$ and $\delta n \leqq x \leqq \nu - \nu^{\frac{1}{3}}$. Then our estimates lead to a term of the form 3A). Let us now examine the case $\delta n \leqq y \leqq \nu - \nu^{\frac{1}{3}}$, $1/n \leqq x \leqq \delta n$. For $\delta n \leqq y \leqq 2\delta n$, we may use estimates on $L_n^{\alpha}(y) - L_{n+2}^{\alpha}(y)$ from row 2 in our tabular estimates to obtain a term of the form 3). If $y \geqq 2\delta n$, $|(x-y)|^{-1} = O(n^{-1})$, and we obtain a term the L^p norm of which is

$$O(1/n)\Big[\int_0^{\delta n} \{x^{-\frac{1}{4}} \int_{2\delta n}^{\nu} f(y)\, (\nu - y)^{\frac{1}{4}}\, dy\}^p\, dx\Big]^{1/p} = O\{\| f \|_p\}$$

by two uses of Holder's inequality. Finally we consider the term

$$\nu - \nu^{\frac{1}{3}} \leqq x \leqq \nu + \nu^{\frac{1}{3}}, \quad \delta n \leqq y \leqq \nu - \nu^{\frac{1}{3}}$$

By using our tabular estimates, we see it suffices to discuss

$$I_1 = n^{-(1/12)} \int_{\delta n}^{\nu - \nu^{\frac{1}{3}}} \frac{f(y)}{x-y}\, (\nu - y)^{\frac{1}{4}}\, dy.$$

But $\dfrac{1}{(x-y)} = \dfrac{1}{\nu - y} + \dfrac{\nu - x}{(\nu - y)(x-y)}$. So $I_1 = I_2 + I_3$.

$$\Big[\int_{\nu - \nu^{(\frac{1}{3})}}^{\nu + \nu^{(\frac{1}{3})}} |I_2(x)|^p\, dx\Big]^{1/p} = O(1)$$

by two uses of Holder's inequality. I_3 reduces to a Hilbert transform.

The proof of Theorem 2′ is the same as the proof of Theorem 1′.

The procedure of deducing Theorems 1 and 2 from 1′ and 2′ is standard provided that Theorems 1 and 2 hold for a set of f dense on L^p (see [9]). The functions $x^n e^{-x/2}$ are dense in $L^p[0,\infty]$. Now if α is an even integer $x^n e^{-(x/2)}$ has a finite expansion in terms of $\mathcal{L}_m^\alpha$, and one obviously has mean convergence. Suppose α is not an even integer. $x^n e^{-(x/2)} = x^{\alpha/2} x^{n-(\alpha/2)} e^{-(x/2)}$, and

$$x^{n-(\alpha/2)} = \Gamma(\alpha/2+n+1) \sum_{m=0}^{\infty} L_m^\alpha(x) \cdot \frac{\Gamma(m-n+\alpha/2)}{\Gamma(\alpha+m+1)}$$

for non-negative integers n, [1], p. 214, equation (16). Therefore

$$x^n = \sum_{m=0}^{\infty} \frac{1}{(m+1)^{n+1}} \mathcal{L}_{m+1}^\alpha(x) + O[\sum_{m=0}^{\infty} \frac{|\mathcal{L}_m^\alpha(x)|}{(m+1)^{n+2}}],$$

and

$$\begin{aligned} x^n - S_j(x) &= \sum_{m=j+1}^{\infty} \frac{1}{(m+1)^{n+1}} \mathcal{L}_m^\alpha(x) \\ &\quad + O[\sum_{m=j+1}^{\infty} \frac{1}{(m+1)^{n+2}} |\mathcal{L}_m^\alpha(x)|] \\ &= \sum_1 + \sum_2. \end{aligned}$$

$\sum_1$ tends to zero as $j \to \infty$ by our estimates for the L^p norms of $\mathcal{L}_m^\alpha(x)$ in Section 1. The same is true for $\sum_2$ unless $n = 0$. Therefore we need only discuss

$$T_j(x) = \sum_{m=j+1}^{\infty} \frac{1}{m+1} \mathcal{L}_m^\alpha(x).$$

From our estimates on Laguerre polynomials, we see

$$\left[\int_0^1 |\mathcal{L}_m^\alpha(x)|^p \, dx\right]^{1/p} \leqq O(m^{-\frac{1}{4}}).$$

Hence we need consider only values of x larger than one. In equation 1, we divide both sides by $y^{\alpha/2}$, and then set $y = 0$. $L_n^\alpha(0) = \dfrac{\Gamma(\alpha+1+n)}{\Gamma(n+1)\Gamma(\alpha+1)}$, [1], p. 189, equation (13). So

$$\begin{aligned} g_n(x) &- \sum_{m=0}^{n} \mathcal{L}_m^\alpha(x) \left\{\frac{\Gamma(\alpha+m+1)}{\Gamma(m+1)}\right\}^{(1/2)} \\ &= O\left\{\frac{1}{x} n^{(1+\alpha/2)}(|\mathcal{L}_n^\alpha(x)| + |\mathcal{L}_{n+1}^\alpha(x)|)\right\} \end{aligned}$$

for $x \geqq 1$. Now we can see from our estimates

$$\left[\int_1^\infty \left| \frac{1}{x} \mathcal{L}_m{}^\alpha(x) \right|^p dx\right]^{1/p} = O[n^{-1/4}] + O[n^{-1/2+1/p}]$$
$$= O(n^{1/4-\xi})$$

for $4/3 < p < 4$, $\xi = \xi(p) > 0$. Thus

$$\left[\int_1^\infty | g_n(x) |^p dx\right]^{1/p} = O\{n^{(\alpha/2+1/2-\xi)}\}.$$

$$\sum_{m=j+1}^\infty \frac{\mathcal{L}_m{}^\alpha(x)}{m+1} = \sum_{m=j+1}^\infty \frac{1}{(m+1)} \left\{ \frac{\Gamma(\alpha+m+1)}{\Gamma(m+1)} \right\}^{-1/2} \cdot (g_m(x) - g_{m-1}(x)).$$

Summing by parts and using Minkowski's inequality, we see

$$\left[\int_1^\infty | T_j(x) |^p dx\right]^{1/p} = O(\sum_{m=j}^\infty m^{-1-\xi}).$$

Thus $\| T_j \|_p$ tends to zero as j tends to infinity.

The proof that Theorem 2′ implies Theorem 2 is similar, but much simpler. Here we take as a dense set the functions $x^n \exp(-x^2/2)$. We use the formulas $\exp(-x^2/2)x^n = \sum_m a_{m,n} \mathcal{H}_m(x)$ where the sum in question is finite for each n. [1], p. 216, equations 28 and 29.

In order to prove the conclusion of Theorem 1′ is false for $1 \leqq p \leqq 4/3$ and $p \geqq 4$, it suffices to prove that $\| \mathcal{L}_n{}^\alpha \|_p \| \mathcal{L}_n{}^\alpha \|_q$ is not bounded for those ranges of p. $(1/p + 1/q = 1)$. For suppose $\| S_n \|_p \leqq M \| f \|_p$ for all f in L^p and all n. Then for fixed n, $\| a_n \mathcal{L}_n{}^\alpha(x) \|_p = \| S_n - S_{n-1} \|_p \leqq 2M$. $\| a_n \mathcal{L}_n{}^\alpha(x) \|_p = | \int f(y) \mathcal{L}_n{}^\alpha(y) dy | \| \mathcal{L}_n{}^\alpha(x) \|_p$. Now there exists an f_n in L^p such that $| \int f_n(y) \mathcal{L}_n{}^\alpha(y) dy | \geqq \frac{1}{2} \| f_n \|_p \| \mathcal{L}_n{}^\alpha \|_q$. Combining these inequalities we find $\| \mathcal{L}_n{}^\alpha(x) \|_p \| \mathcal{L}_n{}^\alpha(x) \|_q \leqq 4M$.

We turn now to the estimate of $\| \mathcal{L}_n{}^\alpha(x) \|_p$ from below. In the following discussion ϵ will be a sufficiently small positive quantity, the exact value of which will vary from place to place. Suppose first $p \leqq 4$. We use Theorem A and the fact that $[\psi'(t)/\psi(t)] = [2t]^{-1}[1 + O(t)]$ to obtain

$$| \mathcal{L}_n{}^\alpha(x) | \geqq (1/\epsilon) | J_\alpha(\nu\psi(x/\nu)) | + O(1/n).$$

This is valid for $1 \leqq x \leqq \epsilon n$. Also $\psi(t) = (t)^{1/2}[1 + O(t)]$. So $\nu\psi(x/\nu) \geqq \epsilon(x\nu)^{1/2}$ for $1 \leqq x \leqq \epsilon n$. Therefore $| \nu\psi(x/\nu) | \geqq \epsilon$ for $1 \leqq x \leqq \epsilon n$. Now for $u \geqq \epsilon$, $| J_\alpha(u) | \geqq \epsilon(u)^{-1/2} | \cos(u + \gamma_\alpha) | + O(u^{-3/2})$; γ_α is some fixed

number. Thus $|\mathcal{L}_n^\alpha(x)| \geqq \epsilon(\nu x)^{-1/4}|\cos(\nu\psi(x/\nu)+\gamma_\alpha)| + O[(\nu x)^{-3/4}+\nu^{-1}]$; $1 \leqq x \leqq \epsilon n$. So for integers l and x such that

$$4_l) \qquad l\pi - \pi/4 \leqq \nu\psi(x/\nu) + \gamma_\alpha \leqq l\pi + \pi/4$$
$$|\mathcal{L}_n^\alpha(x)| \geqq \epsilon(\nu x)^{-(1/4)}.$$

We therefore must see for what values of l and x 4_l holds. Since $\epsilon t^{1/2} \leqq \psi(t) \leqq \epsilon^{-1}t^{1/2}$, 4_l holds for some x if $(1/\epsilon)\nu^{1/2} \leqq l \leqq \epsilon\nu$. Also for a given l, the values of x for which 4_l holds are between $\epsilon l^2/\nu$ and $l^2/(\epsilon\nu)$. This implies the range of values for which 4_l holds is greater than $(\epsilon l)/\nu$. For

$$|\nu\psi(b/\nu) - \nu\psi(a/\nu)| \leqq \int_a^b \psi'(x/\nu)\,dx \leqq 2\int_a^b \frac{\nu^{1/2}}{x^{1/2}}\,dx$$
$$= 4\nu^{1/2}(b^{1/2} - a^{1/2}).$$

So 4_l holds if

$$(b-a) \geqq \epsilon(b^{1/2} + a^{1/2})\nu^{-(1/2)}$$

and if $|\nu\psi(a/\nu) + \gamma_\alpha - \pi l| \leqq \pi/8$.

So 4_l holds in a range of x greater than $(\epsilon l)/\nu$. Finally $|\mathcal{L}_n^\alpha(x)| \geqq \epsilon(l^{-1/2})$ for x satisfying 4_l. So

$$[\int |\mathcal{L}_n^\alpha(x)|^p\,dx]^{1/p} \geqq \epsilon\{\sum_{l=[(1/\epsilon)n^{\frac{1}{2}}]}^{[\epsilon n]} l^{-p/2}(l/\nu)\}^{1/p}$$
$$\geqq \begin{cases} n^{-1/4}(\log n)^{1/4} & p = 4 \\ n^{-1/2+1/p} & p \leqq 4. \end{cases}$$

The analysis of the case $p > 4$ is much simpler because it suffices to consider $\nu - \epsilon\nu^{1/3} \leqq x \leqq \nu + \epsilon\nu^{1/3}$, and in that interval $|Ai(x)|$ is bounded away from 0. One obtains $\| \mathcal{L}_n^\alpha \|_p \geqq A_p n^{-(1/3)+(1/3p)}$.

The proof of the corresponding fact for Hermite polynomials is similar. We just mention that one considers the interval $\delta N^{1/2} \leqq x \leqq N^{1/2} - N^{-(1/6)}$ for $p \leqq 4$ and $N^{1/2} - N^{-(1/6)} \leqq x \leqq N^{1/2} + N^{-1/6}$ for $p > 4$ in estimating the lower bounds of $\| \mathcal{H}_n \|_p$.

UNIVERSITY OF WISCONSIN,
CORNELL UNIVERSITY.

REFERENCES.

[1] A. Erdélyi et al., *Higher Transcendental Functions*, Vol. 2, New York, 1953.

[2] ———, "Asymptotic forms for Laguerre polynomials," *Golden Jubilee Volume of the Indian Mathematical Society*.

[3] ———, "Asymptotic solutions of differential equations with transition points or singularities," *Journal of Mathematical Physics*, vol. 1 (1960), pp. 16-26.

[4] H. and B. S. Jeffreys, *Methods of Mathematical Physics*, 3rd ed., Cambridge, 1956.

[5] G. H. Hardy and J. E. Littlewood, "Some Theorems on Fourier Series and Fourier Power Series," *Duke Mathematical Journal*, vol. 2 (1936), pp. 354-81.

[6] H. Pollard, "The mean convergence of orthogonal Series I," *Transactions of the American Mathematical Society*, vol. 62 (1947), pp. 387-403.

[7] ———, "The mean convergence of orthogonal series II," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 355-367.

[8] G. Szegö, *Orthogonal Polynomials*, 2nd ed., New York, 1959.

[9] G. M. Wing, "The mean convergence of orthogonal series," *American Journal of Mathematics*, vol. 72 (1950), pp. 792-807.

[10] A. Zygmund, *Trigonometric Series*, Vol. I, Cambridge (1959).

A CONCRETE SPECTRAL THEORY FOR SELF-ADJOINT TOEPLITZ OPERATORS.

By Marvin Rosenblum.[1]

1. Introduction. Let $\mathfrak{H}$ be a separable complex Hilbert space and $\mathcal{J}$ a possibly unbounded self-adjoint operator on H. The following formulation of the von Neumann spectral theorem (see [2], chapitre II) describe a diagonalization of $\mathcal{J}$:

There exists a measure σ on the real line R, a σ-measurable function n on R to the non-negative integers or ∞, and an isometric mapping $\mathcal{U}$ on $\mathfrak{H}$ onto a continuous direct sum $\int \mathfrak{H}_\lambda \, d\sigma(\lambda)$ of Hilbert spaces $\mathfrak{H}_\lambda$, $\lambda \in R$, such that for every bounded Baire function β and x, y in $\mathfrak{H}$

1.1 $$\langle \beta(\mathcal{J})x, y\rangle = \int \beta(\lambda) \langle \mathcal{U}x(\lambda), \mathcal{U}y(\lambda)\rangle_\lambda d\sigma(\lambda).$$

$\langle\cdot,\cdot\rangle_\lambda$ is the inner product in $\mathfrak{H}_\lambda$, and the dimension of $\mathfrak{H}_\lambda$ is $n(\lambda)$.

σ and n determine $\mathcal{J}$ up to unitary equivalence; if a self-adjoint operator $\mathcal{J}'$ on $\mathfrak{H}$ has an associated measure σ' and function n' as above, then $\mathcal{J}$ and $\mathcal{J}'$ are unitarily equivalent if and only if σ and σ' are mutually absolutely continuous and n and n' are σ-almost everywhere equal. n is the *multiplicity function* of $\mathcal{J}$, and is determined σ-almost everywhere.

One provides a *spectral theory* for $\mathcal{J}$ when one specifies σ and n. By a *concrete spectral theory* for $\mathcal{J}$ we mean a spectral theory for $\mathcal{J}$ together with a specification of the mapping $\mathcal{U}$.

Let U be the unit circle $\{e^{i\phi}: -\pi < \phi \leqq \pi\}$, and suppose w is a real Lebesgue integrable function on U that is essentially bounded below. As is usual define the Hardy space $\mathfrak{H}^2$ to be the closed span of $\{e^{in\phi}\}_{n=0}^{\infty}$ in $L^2(d\phi)$, where $d\phi$ refers to Lebesgue measure on U. By the *Toeplitz operator $\mathcal{J}$ associated* with w we mean the Friedrichs extension of the semibounded operator $\mathcal{J}_0$ defined on finite linear combinations of $\{e^{in\phi}\}_{n=0}^{\infty}$ by

$$\langle \mathcal{J}_0 f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{i\phi}) g^*(e^{i\phi}) w(e^{i\phi})\, d\phi.$$

$*$ denotes the complex conjugation operation, defined so $g^*(e^{i\phi}) = (g(e^{i\phi}))^*$.

Received August 14, 1964.

[1] This research was supported by National Science Foundation Grant 18853.

The matrix representation of the self-adjoint operator $\mathcal{T}$ with respect to the complete orthonormal set $\{e^{in\phi}\}_{n=0}^{\infty}$ in $\mathfrak{H}^2$ is $[w_{j-k}]_{j,k=0}$ where

$$w_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} w(e^{i\phi}) e^{-ij\phi} d\phi, \qquad j = 0, \pm 1, \pm 2, \cdots.$$

This matrix is a Toeplitz matrix of the type studied in [3].

We will be concerned in this note with providing a concrete spectral theory for Toeplitz operators $\mathcal{T}$ associated with w, and we shall sketch how to apply our theory to a class of self-adjoint Wiener-Hopf operators. We demand throughout that w is not almost everywhere equal to a constant function. This will eliminate the trivial situations where $\mathcal{T}$ is a scalar multiple of the identity. We continue work started by Hartman and Wintner in [4] and [5]. From their analysis it follows that σ has no atoms and the support of σ is the closed interval $[\operatorname{ess\,inf} w, \operatorname{ess\,sup} w]$. Putnam [8] proved under subsidiary conditions on w that σ is absolutely continuous with respect to Lebesgue measure. In [9] I showed that σ always enjoys that property. Thus we can and will take σ to be Lebesgue measure on $[\operatorname{ess\,inf} w, \operatorname{ess\,sup} w]$. In [10] I provided a concrete spectral theory for $\mathcal{T}$ under the assumption that for each real λ the set $\Gamma_\lambda = \{e^{i\phi} : w(e^{i\phi}) \leqq \lambda\}$ is, modulo a set of measure zero, an arc of the circle or null. The result in this paper implies that these operators are precisely those $\mathcal{T}$ of multiplicity one.

Ismagilov, in a Doklady note [7], indicated proofs of generalizations of the above results, and described the multiplicity function n, thus completing a (nonconcrete) spectral theory for $\mathcal{T}$. We shall codify Ismagilov's description of n, which we discovered independently and find to be of independent interest. We shall introduce a concept of index or winding number for measurable real-valued functions f on U, in terms of which n can be very simply described; $n(\lambda)$ for $\mathcal{T}$ associated with w turns out to be the index of $\lambda - w$.

2. Index of a real function on U. Suppose f is any real Lebesgue measurable function on U. Let $\Gamma = \{e^{i\phi} : f(e^{i\phi}) \geqq 0\}$. Necessarily Γ satisfies exactly one of the following alternatives almost everywhere:

$A(0)$, Γ is empty or $\Gamma = U$; or

$A(n)$, Γ is the union of n disjoint closed non-degenerate arcs, where n is a positive integer; or

$A(\infty)$, neither of the above hold.

We define the *index* or *winding number* $I(f)$ of f (with respect to the origin) to be x in case $A(x)$ holds.

We shall provide some algebraic implementation to the geometric idea of index in Theorem 1. First we associate with f a holomorphic function F and a measure μ on U. Define F by

$$2.1 \qquad i\pi F(z) = \exp[\tfrac{1}{2}i \int_\Gamma k(ze^{i\phi})\, d\phi] - \cos[\tfrac{1}{2} \int_\Gamma d\phi], \; |z| < 1,$$

where $k(u) = (1+u)(1-u)^{-1}$. The additive constant in 2.1 normalizes F so that $\operatorname{Im} F(0) = 0$. F has non-negative real part in the unit disk, so there exists a measure μ on U such that

$$2.2 \qquad F(z) = \int k(ze^{i\alpha})\, d\mu(\alpha).$$

Put $c_j = \int e^{-ij\alpha}\, d\mu(\alpha)$ for all integers j, and define the $(m+1) \times (m+1)$ Toeplitz matrix M_m by specifying that the (j,k) entry is c_{j-k}, $j, k = 0, 1, \cdots, m$.

LEMMA 1. *μ is a singular measure.*

Proof. Suppose $re^{i\theta}$ is the polar representation of z, $|z| < 1$. $\operatorname{Re} k(z)$ and $\operatorname{Im} k(z)$ are the Poisson and conjugate Poisson kernels, with

$$\operatorname{Re} k(z) = (1-r^2)(1-2r\cos\theta + r^2)^{-1}, \text{ and}$$
$$\operatorname{Im} k(z) = r\sin\theta(1-2r\cos\theta+r^2)^{-1}.$$

Thus

$$\begin{aligned}\int \operatorname{Re} k(ze^{i\alpha})\, d\mu(\alpha) &= \operatorname{Re} F(z)\\ &= \operatorname{Re}(-i\pi^{-1}\exp[\tfrac{1}{2}i\int_\Gamma k(ze^{i\phi})\, d\phi])\\ &= \pi^{-1}\exp[-\tfrac{1}{2}\int_\Gamma \operatorname{Im} k(ze^{i\phi})\, d\phi]\cdot\\ &\qquad \sin[\tfrac{1}{2}\int_\Gamma \operatorname{Re} k(ze^{i\phi})\, d\phi]\end{aligned}$$

As $r \to 1$ this almost everywhere approaches

$$\pi^{-1}\exp[-\tfrac{\pi}{2}\mathfrak{X}^{\sim}]\sin(\pi\mathfrak{X}),$$

where $\mathfrak{X}$ is the indicator function of Γ and $\mathfrak{X}^{\sim}$ is the Fourier conjugate of $\mathfrak{X}$. Thus a. e.

$$\lim_{r\to 1} \int \operatorname{Re} k(ze^{i\alpha})\, d\mu(\alpha) = 0, \text{ so}$$

μ is a singular measure. See [1], p. 386, for a related result.

LEMMA 2. *Suppose n is a positive integer.*

i) *If* Γ *is the union of* n *disjoint arcs*

$$\{e^{i\phi}: a_j \leqq \phi \leqq b_j, j = 1,\cdots,n\},$$

then

2.3 $$i\pi F(z) = \prod_{j=1}^{n} e^{\frac{1}{2}i(b_j-a_j)}(1-ze^{ia_j})(1-ze^{ib_j})^{-1} - \cos[\tfrac{1}{2}\sum_{j=1}^{n}(b_j-a_j)]$$

2.4 $$= i\pi \sum_{j=1}^{n} \rho_j k(e^{ib_j}z).$$

ρ_j *is the residue of* $-\frac{1}{2}e^{ib_j}F$ *at* e^{-ib_j} *and the* ρ_j *are positive,* $j = 1, 2,\cdots,n$.

ii) *Conversely suppose that* F *associated with* Γ *by* 2.1 *is of the form* 2.4 *with* $\rho_1,\cdots,\rho_n$ *positive, and* $e^{ib_1},\cdots,e^{ib_n}$ *are distinct points on* U. *Then* $I(f) = n$.

Proof. If Γ is the union of n arcs as described in i), then 2.3 follows from 2.1 by direct computation. The F defined by 2.3 is rational, holomorphic in the unit disk, satisfies $F^*(z^{-1}) = -F(z^*)$ for all complex z, and has n poles. Therefore by a simple argument (see [3], p. 151), F is of the form 2.4.

Conversely suppose F is a rational function of the form 2.4 as hypothesized in ii). Put

$$G(z) = i\pi F(z) + \cos[\tfrac{1}{2}\int_\Gamma d\phi], |z| < 1.$$

Then

$$\operatorname{Im}\log G(z) = \tfrac{1}{2}\int_\Gamma \operatorname{Re} k(ze^{i\phi})\,d\phi,$$

so $\operatorname{Im}\log G = \pi\mathfrak{X}$ a. e. on U, where $\mathfrak{X}$ is, as before, the indicator function of Γ. Thus we see that G is a. e. real on U, and G is a. e. positive at points of U where $\mathfrak{X} = 0$. Necessarily then $I(f) = n$ with the endpoints of the arcs of Γ being zeros and poles of G. The above argument was employed in [1], p. 331 for the half-line analogues of equations 2.1 and 2.2.

THEOREM 1. i) *The dimension of the complex Hilbert space* $L^2(d\mu)$ *is equal to* $I(f)$.

ii) $I(f)$ *is equal to the smallest non-negative integer* n *such that* $\det M_n = 0$; *or if no such integer exists, then* $I(f) = \infty$.

Proof. It is clear from 2.1 that $I(f) = 0$ if and only if F vanishes identically, that is, in case μ is the null measure. Thus $I(f) = 0$ if and only if the dimension of $L^2(d\mu)$ is 0.

From Lemma 2 ii) it follows that $I(f) = n$, $0 < n < \infty$, if and only

if F is of the form 2.4, and this is the case when μ is purely atomic with exactly n mass points. Thus $I(f)=n$ if and only if $\dim L^2(d\mu)=n$. The remaining case $I(f)=\infty$ can then only occur if and only if $\dim L^2(d\mu)=\infty$. This completes the proof of i).

We next consider ii). Since μ is singular it follows from the F. and M. Riesz theorem that $\{e^{in\alpha}\}_{n=0}^{\infty}$ is total in $L^2(d\mu)$. This is the case since if some g in $L^2(d\mu)$ satisfies

$$\int e^{-in\alpha}g(e^{i\alpha})\,d\mu(\alpha)=0 \text{ for } n=0,1,\cdots,$$

then $g\,d\mu$ is both absolutely continuous and singular, so $g=0$ a.e. Next we apply the Gram-Schmidt process to 1, $e^{i\alpha}, e^{2i\alpha},\cdots$, in $L^2(d\mu)$ and obtain a set of orthonormal polynomials $p_0,\cdots,p_{n-1}$ provided the Gram determinant $\det M_{n-1}$ is not zero (see [3], p. 37). If $\det M_n\neq 0$ for all positive integers n, then the dimension of $L^2(d\mu)$ is ∞, and, by i), $I(f)=\infty$. Next suppose that there is positive integer n such that $\det M_{n-1}\neq 0$ and $\det M_n=0$, so $e^{in\alpha}$ is in the span of $p_0,\cdots,p_{n-1}$ in $L^2(d\mu)$ norm. Then $e^{in\alpha}$ is in the $L^2(d\mu)$ span of $\{e^{ij\alpha}\}_{j=0}^{n-1}$, and hence so is $e^{i(n+1)\alpha}$, $e^{i(n+2)\alpha}$, etc. Thus $p_0,\cdots,p_{n-1}$ is a complete orthonormal set in $L^2(d\mu)$ and $L^2(d\mu)$ has dimension n. Thus $I(f)=n$. Finally, $\det M_0=c_0=0$ if and only if μ is the null measure, which is the case only if $I(f)=0$.

3. A concrete spectral theory for $\mathcal{T}$. We shall build on the self-adjoint Toeplitz operator theory developed in [9] and [10]. Suppose that u and v are complex numbers of modulus less than one. Define k_u and k_v in $\mathfrak{H}^2$ so

3.1 $$\langle \mathcal{T}k_u,k_v\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}(1-ue^{i\phi})^{-1}(1-v^*e^{-i\phi})^{-1}w(e^{i\phi})\,d\phi.$$

For each real λ set

3.2 $$\Gamma_\lambda=\{e^{i\phi}:\lambda-w(e^{i\phi})\geqq 0\},$$

and let F_λ be the holomorphic function and μ_λ the measure associated with $\lambda-w$ as in Section 2. Let S be a subset of the real line R such that $\log|w-\lambda|$ is in $L^1(d\phi)$ for all λ in S and $R\setminus S$ has measure zero. Such an S exists by [10], Lemma 1. Define for each λ in S the zero-free function $\xi(\cdot,\lambda)$ on the unit disk by

3.3 $$\begin{aligned}\xi(u,\lambda)&=\lim_{\epsilon\downarrow 0}\exp\{-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log(w(e^{i\phi})-\lambda+i\epsilon)k(ue^{i\phi})\,d\phi\}\\&=\exp\{-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log|w(e^{i\phi})-\lambda|\,k(ue^{i\phi})\,d\phi\}\cdot\\&\qquad\exp[-\frac{i}{4}\int_{\Gamma_\lambda}k(ue^{i\phi})\,d\phi].\end{aligned}$$

We reformulate a pertinent result from [9] and [10] in

THEOREM 2. *Let β be any bounded complex Baire function on R. Then*

3.4 $$\langle \beta(\mathcal{J}) k_u, k_v \rangle = \int_{-\infty}^{\infty} \beta(\lambda) \int \frac{\xi(u,\lambda)}{1-ue^{i\alpha}} \left(\frac{\xi(v,\lambda)}{1-ve^{i\alpha}} \right)^* d\mu_\lambda(\alpha)\, d\lambda.$$

Proof. From [10] we have

$$\langle \beta(\mathcal{J}) k_u, k_v \rangle = \int_{-\infty}^{\infty} \beta(\lambda)\, d\langle E(\lambda) k_u, k_v \rangle$$

$$= \tfrac{1}{2} \int_{-\infty}^{\infty} \beta(\lambda) \xi(u,\lambda) \xi^*(v,\lambda) (1-uv^*)^{-1} (F_\lambda(u) + F_\lambda^*(v))\, d\lambda,$$

which, by the definition of μ_λ, equals the right side of 3.4.

Next we shall employ the theory of continuous direct sums of Hilbert spaces as set forth in [2], chapitre II, to recast 3.4 in the form 1.1. We use the notation $\langle \cdot, \cdot \rangle_\lambda$ for the inner product in $L^2(d\mu_\lambda)$, $\lambda \in S$, and consider the vector field $\{k_u : u \in D\}$, where D is a countable dense subset of the open unit disk in the complex plane. We define the operation $\mathcal{U}$ by

3.5 $$\mathcal{U} k_u(\lambda) = (1-ue^{i\alpha})^{-1} \xi(u,\lambda),$$

so $\mathcal{U} k_u(\lambda)$ is in $L^2(d\mu_\lambda)$ for each λ in S and u in D. For any bounded Baire function β we can put

$$\mathcal{U}(\beta(\mathcal{J}) k_u)(\lambda) = \beta(\lambda) \mathcal{U} k_u(\lambda).$$

LEMMA 3. i) *The functions $\lambda \to \langle \mathcal{U} k_u(\lambda), \mathcal{U} k_v(\lambda) \rangle_\lambda$ are Lebesgue measurable for each u, v in D.*

ii) *If $\lambda \in S$, then $\{\mathcal{U} k_u(\lambda) : u \in D\}$ is total in $L^2(d\mu_\lambda)$.*

iii) $\| k_u \|^2 = \int_{-\infty}^{\infty} \| \mathcal{U} k_u(\lambda) \|_\lambda^2\, d\lambda$ *for each u in D.*

Proof. i) and iii) follow immediately from Theorem 2, so we turn to the proof of ii). Assume λ is in S and that a vector g in $L^2(d\mu_\lambda)$ satisfies

$$\xi^*(u,\lambda) \int g(e^{i\alpha}) (1-u^* e^{-i\alpha})^{-1}\, d\mu_\lambda(\alpha) = 0$$

for all u in D. Then $\int g(e^{i\alpha}) e^{-i\alpha j}\, d\mu_\lambda(\alpha) = 0$ for $j = 0, 1, \cdots$. Thus by the F. and M. Riesz theorem, [6] p. 47, the measure $g\, d\mu_\lambda$ is absolutely continuous with respect to Lebesgue measure. But by Lemma 1 it is also singular, so μ_λ or g is null. This proves ii).

From [2], Propositions 4, 144 and 7, p. 149, it now follows that a continuous direct sum

$$\mathfrak{D} = \int_{-\infty}^{\infty} L^2(d\mu_\lambda)\, d\lambda \text{ exists and}$$

$$\{\mathcal{U}(\beta(\mathcal{J})k_u)(\lambda) : \beta \text{ is a bounded Baire function and } u \in D\}$$

is a total set in $\mathfrak{D}$. From this we deduce that $\mathcal{U}$ has a unique linear extension to an isometric mapping on H^2 onto $\mathfrak{D}$. We can now formulate our main result.

THEOREM 3. *Suppose w is a real $L^1(d\phi)$ function that is bounded below and that $\mathcal{J}$ is the self-adjoint Toeplitz operator associated with w. For each real λ define Γ_λ by 3.2, and the measure μ_λ by 2.1 and 2.2 with "Γ" replaced by "Γ_λ."*

Then there exists an isometric mapping $\mathcal{U}$ of H^2 onto a continuous direct sum $\int_{-\infty}^{\infty} L^2(d\mu_\lambda)\,d\lambda$ such that 1.1 holds with σ equal to Lebesgue measure. The multiplicity function n of $\mathcal{J}$ is for each real λ equal to the dimension $n(\lambda) = I(\lambda - w)$ of $L^2(d\mu_\lambda)$. $\mathcal{U}$ is defined on

$$\{k_u : |u| < 1, k_u(e^{i\phi}) = (1 - ue^{i\phi})^{-1}\}$$

by 3.5.

We next set down two variations on Theorem 3 that may be more useful for some applications. Let $l_{2,n(\lambda)}$ be $n(\lambda)$-dimensional sequential Hilbert space. For each λ in S let $p_0(e^{i\alpha},\lambda),\cdots p_{n(\lambda)-1}(e^{i\alpha},\lambda)$ be the orthonormal polynomials obtained from applying the Gram-Schmidt process to $1, e^{i\alpha}, e^{2i\alpha},\cdots,$ in $L^2(d\mu_\lambda)$ as described in the proof of Theorem 1. Define $\eta_j(\cdot,\lambda)$ in the unit disk by

$$\eta_j(u,\lambda) = \int (1 - ue^{i\alpha})^{-1} p^*_{j-1}(e^{i\alpha},\lambda)\, d\mu_\lambda(\alpha), \qquad j = 1,\cdots, n(\lambda),$$

and set $\eta_j(u,\lambda) = 0$ if λ is not in $[\operatorname{ess\,inf} w, \operatorname{ess\,sup} w]$.

COROLLARY 1. *Assume the hypotheses of Theorem 3. Then there exists an isometric mapping $\mathcal{V}$ of $\mathfrak{H}^2$ onto a continuous direct sum $\int_{-\infty}^{\infty} l_{2,n(\lambda)}\,d\lambda$ such that for all bounded Baire functions β and all x, y in $\mathfrak{H}^2$*

$$\langle \beta(\mathcal{J})x, y\rangle = \int_{-\infty}^{\infty} \beta(\lambda) \sum_{j=1}^{n(\lambda)} (\mathcal{V}x)_j(\lambda)\,(\mathcal{V}y)_j{}^*(\lambda)\, d\lambda. \tag{3.6}$$

$\mathcal{V}$ *is defined on the* k_u *by*

$$(\mathcal{V}k_u)_j(\lambda) = \xi(u,\lambda)\eta_j(u,\lambda), \qquad j=1,\cdots,n(\lambda).$$

Proof. Suppose $\lambda \in S$. Then

$$\langle \mathcal{U}k_u(\lambda), \mathcal{U}k_v(\lambda)\rangle_\lambda = \xi(u,\lambda)\xi^*(v,\lambda)\int(1-ue^{i\alpha})^{-1}(1-v^*e^{-i\alpha})^{-1}d\mu_\lambda(\alpha),$$

which by the Parseval equality equals

$$\sum_{j=1}^{n(\lambda)}(\mathcal{V}x)_j(\lambda)(\mathcal{V}y)_j^*(\lambda).$$

Thus Corollary 1 is now a direct consequence of Theorem 3.

COROLLARY 2. *Suppose* $\mathcal{T}$ *is as in Theorem 3 and further suppose that* n *is almost everywhere finite, so* Γ_λ *can be taken to be the union of* $n(\lambda)$ *disjoint arcs*

$$\{e^{i\phi}\colon a_j(\lambda) \leqq \phi \leqq b_j(\lambda), j=1,\cdots,n(\lambda)\}$$

for each λ *in* $[\operatorname{ess\,inf} w, \operatorname{ess\,sup} w]$. *For any other* λ *set* $a_j(\lambda) \equiv b_j(\lambda) \bmod 2\pi$. *Put* $\rho_j(\lambda)$ *equal to* $-\frac{1}{2}e^{ib_j(\lambda)}$ *times the residue of* F_λ *at* $e^{-ib_j(\lambda)}$.

There exists an isometric mapping $\mathcal{W}$ *of* $\mathfrak{H}^2$ *onto* $\int_{-\infty}^{\infty} l_{2,n(\lambda)}\,d\lambda$ *such that* 3.6 *holds with "*$\mathcal{V}$*" replaced by "*$\mathcal{W}$*."* $\mathcal{W}$ *is defined on the* k_u *by*

$$(\mathcal{W}k_u)_j(\lambda) = (\rho_j(\lambda))^{\frac{1}{2}}(1-ue^{ib_j(\lambda)})^{-1}\xi(u,\lambda), \qquad j=1,\cdots,n(\lambda).$$

Proof. It follows from equations 2.3 and 2.4 in Lemma 2 that μ_λ consists of $n(\lambda)$ mass points at $\{e^{ib_j(\lambda)}\colon j=1,\cdots,n(\lambda)\}$ and there is a mass of $\rho_j(\lambda)$ at $e^{ib_j(\lambda)}$. Thus

$$\langle \mathcal{U}k_u(\lambda), \mathcal{U}k_v(\lambda)\rangle_\lambda = \sum_{j=1}^{n(\lambda)}\mathcal{W}k_u(\lambda)(\mathcal{W}k_v)^*(\lambda).$$

4. Examples and applications.

i) The following representation for $\mathfrak{H}^2$ functions x follows directly from Theorem 3:

$$x(v^*) = \int_{-\infty}^{\infty}\langle \mathcal{U}x(\lambda), \mathcal{U}k_v(\lambda)\rangle_\lambda\, d\lambda, \ |v|<1.$$

If in addition x is in the domain of $\mathcal{T}$, then

$$(\mathcal{T}x)(v^*) = \int_{-\infty}^{\infty}\lambda\langle \mathcal{U}x(\lambda), \mathcal{U}k_v(\lambda)\rangle_\lambda\, d\lambda.$$

ii) A Toeplitz operator has spectrum of multiplicity one if and only if Γ_λ is an arc of U for each λ in $(\operatorname{ess\,inf} w, \operatorname{ess\,sup} w)$. Suppose

$$\Gamma_\lambda = \{e^{i\phi}\colon a(\lambda) \leqq \phi \leqq b(\lambda)\}$$

for each real λ. We employ Corollary 2, and find

$$i\pi F_\lambda(z) = e^{\frac{1}{2}i(b(\lambda)-a(\lambda))}(1 - ze^{ia(\lambda)})\,(1 - ze^{ib(\lambda)})^{-1} - \cos[\tfrac{1}{2}(b(\lambda) - a(\lambda))],$$

$\rho(\lambda) = \pi^{-1}\sin[\frac{1}{2}(b(\lambda) - a(\lambda)]$. Thus for all bounded Baire functions β and x, y in $\mathfrak{H}^2$

$$\langle\beta(T)x, y)\rangle = \int_{-\infty}^{\infty} \beta(\lambda)\,\mathfrak{V}x(\lambda)\,(\mathfrak{V}y)^*(\lambda)\,d\lambda,$$

where $\mathfrak{V}k_u(\lambda) = \rho^{\frac{1}{2}}(\lambda)\,(1 - ue^{ib(\lambda)})^{-1}\xi(u,\lambda)$. This is, modulo a change of notation, the diagonalization derived in [10].

iii) We give an example to illustrate Corollary 1. Suppose w, a, b, ξ, ρ, are as in ii), and suppose that n is a positive integer. We consider the Toeplitz operator $\mathcal{T}_n$ associated with $w(e^{i\phi n})$. Γ_λ consists of n arcs for each λ in $(\operatorname{ess\,inf} w, \operatorname{ess\,sup} w)$, so $\mathcal{T}_n$ has spectrum of multiplicity n. One finds that the measure μ_λ has n mass points at

$$e^{ib(\lambda)/n}e^{2\pi ij/n}, \qquad j = 1, \cdots, n$$

and each point has mass $\rho(\lambda)/n$. Further computations yield

$$p_j(e^{i\alpha}, \lambda) = \rho^{-\frac{1}{2}}(\lambda)\,e^{i\alpha j},$$

and

$$(\mathfrak{V}k_u)_j = \rho^{\frac{1}{2}}(\lambda)\,u^{j-1}(1 - u^n e^{ib(\lambda)})^{-1}\xi(u^n, \lambda) \qquad j = 1, \cdots, n.$$

iv) Consider the self-adjoint integral operator $\mathcal{K}$ on $L^2(0,\infty)$ defined by

$$(\mathcal{K}f)(x) = \int_0^\infty K(x-y)f(y)\,dy,$$

where $K \in L^1(-\infty, \infty)$ and $K^*(x) = K(-x)$ for all x. $\mathcal{K}$ is a *Wiener-Hopf* integral operator. Define the Laguerre functions $\{\Phi_n\}_0^\infty$ on $[0,\infty)$ by the generating function formula

$$(1-u)^{-1}\exp[-\tfrac{1}{2}k(u)x] = \sum_0^\infty \Phi_n(x)u^n, \ |u| < 1.$$

It turns out that the matrix representation of $\mathcal{K}$ with respect to complete orthonormal set $\{\Phi_n\}_0^\infty$ is a Toeplitz matrix. Thus the theory developed in

this paper can be applied to diagonalize $\mathcal{K}$. We shall present the details in a subsequent paper.

UNIVERSITY OF VIRGINIA.

REFERENCES.

[1] N. Aronszajn and W. F. Donoghue, Jr., "On exponential representations of analytic functions in the upper half-plane with positive imaginary part," *Journal d'Analyse Mathématique*, vol. 5 (1956-57), pp. 321-388.

[2] J. Dixmier, *Les algèbres d'opérateurs dans l'espace hilbertien* (*Algebres de von Neumann*), Cahiers scientifiques, Fas. XXV, Gauthier-Villars, Paris (1957).

[3] U. Grenander and G. Szegö, *Toeplitz forms and their applications*, University of California, 1958.

[4] P. Hartman and A. Wintner, "On the spectra of Toeplitz's matrices," *American Journal of Mathematics*, vol. 72 (1950), pp. 359-366.

[5] ———, "The spectra of Toeplitz's matrices," *American Journal of Mathematics*, vol. 76 (1954), pp. 867-882.

[6] K. Hoffman, *Banach spaces of analytic functions*, Englewood Cliffs, N. J. (1962).

[7] R. S. Ismagilov, "The spectrum of Toeplitz matrices," *Doklady Akademii Nauk SSSR*, vol. 149 (1963), pp. 769-772. Translated in *Soviet Mathematics*, vol. 4 (1963), pp. 462-465.

[8] C. R. Putnam, "On Toeplitz matrices, absolute continuity, and unitary equivalence," *Pacific Journal of Mathematics*, vol. 9 (1959), pp. 837-846.

[9] M. Rosenblum, "The absolute continuity of Toeplitz's matrices," *Pacific Journal of Mathematics*, vol. 10 (1960), pp. 987-996.

[10] ———, "Self-adjoint Toeplitz operators and associated orthonormal functions," *Proceedings of the American Mathematical Society*, vol. 13 (1962), pp. 590-595.

ALMOST AUTOMORPHIC FUNCTIONS ON GROUPS.

By W. A. Veech.

Introduction. S. Bochner has observed in various contexts that a certain property enjoyed by the almost periodic functions on a group G can be used with advantage in obtaining simpler and conceptually more natural proofs of certain theorems concerning these functions. ([2], [4], [5].) Bochner calls his property "almost automorphy" because it first arose in work on differential geometry. Taking G for the present to be the group of integers $(= \boldsymbol{Z})$ an almost automorphic function f has the property that from any sequence $\{\alpha_n'\} \subset \boldsymbol{Z}$ may be extracted a subsequence $\{\alpha_n\}$ such that both $\lim_{n\to\infty} f(t+\alpha_n) = g(t)$ and $\lim_{n\to\infty} g(t-\alpha_n) = f(t)$ hold for each $t \in \boldsymbol{Z}$ and some function g, but *not necessarily uniformly*. Bochner has observed that almost periodic functions are almost automorphic, but the converse is not true. ([5], [18].) However we will show in the present paper that the almost automorphic functions on a group can be characterized in terms of the almost periodic functions. A function f on G is almost automorphic if and only if it is the pointwise limit of a "jointly almost automorphic" net of almost periodic functions. (A consequence of this result is that a group is maximally (minimally) almost automorphic if and only if it is maximally (minimally) almost periodic.) Conversely one can characterize almost periodicity in terms of almost automorphy: A function f is almost periodic if and only if $\lim_n f(t+\alpha_n) = g(t)$ is almost automorphic whenever the limit exists. This latter result is shown in Section 5 to be applicable to classical theorems of Favard [8] concerning almost periodic solutions of differential equations, the relevance of almost automorphy to which was shown by Bochner. ([5].)

Essential for our discussion of almost automorphic functions has been the notion of a Bohr almost automorphic function which is introduced in Section 2. Briefly stated for $G = \boldsymbol{Z}$ we require of a Bohr almost automorphic function f that it be bounded, and that for each $s \in \boldsymbol{Z}$ and $\epsilon > 0$ there is a relatively dense subset $B_\epsilon = B_\epsilon(s)$ of $\boldsymbol{Z}$ such that if $\tau_1, \tau_2 \in B_\epsilon$, then $|f(s+\tau_1-\tau_2) - f(s)| < 2\epsilon$. A similar condition for $G = \boldsymbol{R}$ was introduced by B. M. Levitan. (We mistakenly stated in [18] that Levitan's class of

Received January 15, 1964.

N-a.p. functions ([12]) is essentially identical with ours. It is now not clear to us what the relationship between the two classes is.) Theorem 2.2.1 asserts the equivalence of Bochner almost automorphy with Bohr almost automorphy.

In Section 3 following a suggestion of Professor Harry Furstenberg we consider the flow associated with an almost automorphic function. For $G=\boldsymbol{Z}$ one considers the closure X of the space of translates of the given function f, the topology being convergence on finite sets. The operations of translation extend to be continuous on X and $\boldsymbol{Z}$ is there represented by homeomorphisms. Thus there is a flow $(\boldsymbol{Z},X)$ associated with the function f. A point $\xi\in X$ is said to be almost automorphic if from any sequence $\{\alpha_n'\}\subset\boldsymbol{Z}$ may be extracted a subsequence $\{\alpha_n\}$ such that $\lim_{m\to\infty}\lim_{n\to\infty}(\alpha_n-\alpha_m)\xi=\xi$. Theorem 3.2.1 answers in the affirmative a question posed to us by Furstenberg. We prove the existence of a compact group X_0 upon which $\boldsymbol{Z}(G)$ acts and a closed mapping $\pi\colon X\to X_0$ such that the diagram

$$\begin{tikzcd} X \arrow[r, "\boldsymbol{Z}"] \arrow[d, "\pi"'] & X \arrow[d, "\pi"] \\ X_0 \arrow[r, "\boldsymbol{Z}"] & X_0 \end{tikzcd}$$

is commutative and $\pi^{-1}\pi x=\{x\}$ exactly when x is an almost automorphic point. (The almost automorphic points of X are dense in X.) The following example due to Furstenberg (oral communication) is illuminating as an illustration of Theorem 3.2.1. Let θ be an irrational real number. For $n\in\boldsymbol{Z}\cos 2\pi n\theta$ is never zero, and we may form the function

$$f(n)=\operatorname{sgn}\cos 2\pi n\theta=\begin{cases}+1 & \cos 2\pi n\theta>0\\ -1 & \cos 2\pi n\theta<0\end{cases}.$$

It can be shown that $f(n)$ is almost automorphic but not almost periodic. A typical translate of f is given by $\operatorname{sgn}\cos 2\pi(n+m)\theta$. The closure X of such translates would ideally be those functions $f_x(n)=\operatorname{sgn}\cos 2\pi(n\theta+x)$ except that for certain (countably many) x, $0\leqq x\leqq 1$ there exists an integer n with $\cos 2\pi(n\theta+x)=0$. For this pair $n,x\operatorname{sgn}\cos 2\pi(n\theta+x)$ is not well defined, and this phenomenon exhibits itself in X by the presence of two points ξ_1 and ξ_2 with $\xi_1=\operatorname{sgn}\cos 2\pi(n\theta+x)$, $\operatorname{sgn}0=+1$, $\xi_2=\operatorname{sgn}\cos 2\pi(n\theta+x)$, $\operatorname{sgn}0=-1$. Thus X appears as a "covering" of the circle $\Gamma=\boldsymbol{R}/\boldsymbol{Z}$ with the property that one point lies above all but countably many points over which lie two points of X.

We make use of the space X_0 in Section 3 to produce a net of almost periodic functions on the group $\boldsymbol{Z}(G)$ which converges pointwise to f and has the required property of joint almost automorphy. Since $\boldsymbol{Z}$ acts on X_0, a continuous function g on X_0 can be considered a function on $\boldsymbol{Z}$ if we define $h(n)=g(n\xi)$ for a fixed point $\xi\in X_0$. Since X_0 is a compact group, $h(n)$ is almost periodic on $\boldsymbol{Z}$ because it is the restriction to $Z\xi\subset X_0$ of the almost periodic function g on X_0. We will not attempt here to describe the procedure for producing the desired net.

In Section 4 we introduce certain countability and continuity assumptions and then show for a class of abelian groups (containing $\boldsymbol{Z}^n, \boldsymbol{R}^n$) that there is a Fourier analysis for the almost automorphic functions. Each such function f possesses a (non-unique) expansion $f(t)\sim\sum_\lambda a_\lambda\chi_\lambda(t)$ where χ_λ is is a continuous character on G. It is shown using a result of Bogolilouboff-Følner that the Fourier series for $f(t)$ can be summed in a "jointly almost automorphic" way to f and hence uniquely determines f.

We conclude in Section 6 with some examples announced earlier ([18].)

The results presented here are an extension of the author's doctoral thesis [17] written at Princeton University under the guidance of Professor S. Bochner. I would like once more to express my warm thanks to Professor Bochner for his kindness in general and invaluable advice in particular during the preparation of both thesis and the present paper. I would also like to thank Professor Harry Furstenberg whose suggestion it was as mentioned before to analyze the flow associated with a function. Before this we had been unable to characterize the almost automorphic functions on non-commutative groups. Finally I would like to thank the National Science Foundation for generous support during the predoctoral phase of this research.

1. Definitions and elementary properties.

1.1. Notation and conventions. Given a complex valued function f on a group G new functions γf, $f\gamma$ are defined for each $\gamma\in G$ by $\gamma f\cdot(t)=f(\gamma t)$, $f\gamma\cdot(t)=f(t\gamma)$. The former is a left translate of f while the latter is a right translate of f. We note that for $\gamma,\delta\in G$ $(\gamma\delta)f=\delta(\gamma f)$ and $(f\delta)\gamma=f\gamma\delta$. Given a net $\{\alpha_i\}_{i\in\Lambda}$ of group elements such that $\lim_{i\in\Lambda}\alpha_i f=g$ exists pointwise on G we shall write $g=T_\alpha f$. If $\lim_{i\in\Lambda}\alpha_i^{-1}g=h$ exists we write $h=T_{\alpha^{-1}}g$. The symbols for the corresponding limits of right translates will be $S_\alpha f$, $S_{\alpha^{-1}}g$. Given nets $\{\alpha_i\}_{i\in\Lambda_1}, \{\beta_j\}_{j\in\Lambda_2},\cdots,\{\gamma_k\}_{k\in\Lambda_n}$ the expression $T_\alpha T_\beta\cdots T_\gamma f$ will be taken to mean $\lim_{i\in\Lambda_1}\lim_{j\in\Lambda_2}\cdots\lim_{k\in\Lambda_n}(\gamma_k\cdots\beta_j\alpha_i)f$ the limits existing point-

wise on G. Finally we will generally denote a subnet of a net given by $\alpha'=\{\alpha'_i\}_{i\in\Lambda'}$ by removing the primes. That is $\alpha=\{\alpha_i\}_{i\in\Lambda}$ means that $\alpha_i=\alpha_{j(i)}'$, $i\in\Lambda$, $j(i)\in\Lambda'$, and the set $\{j(i)\}$ is cofinal in Λ'.

1.2. Bochner almost automorphic functions. We present here what is essentially Bochner's definition of an almost automorphic function on a group. We have deviated from his definition only in the use of nets rather than sequences, an alteration evidently necessary for treating the general case. In Section 4 certain assumptions will be introduced under which sequences suffice.

Definition 1.2.1. A complex valued function on a group G is *left almost automorphic* if any net $\alpha'=\{\alpha_i'\}_{i\in\Lambda'}$ of group elements has a subnet $\alpha=\{\alpha_i\}_{i\in\Lambda}$ such that both

$$T_\alpha f = g \tag{1.2.1}$$

and

$$T_{\alpha^{-1}} g = f \tag{1.2.2}$$

hold on G for some complex valued (i.e. finite) function g. If S_α, $S_{\alpha^{-1}}$ are substituted in the definition for T_α, $T_{\alpha^{-1}}$ respectively, then we say f is *right almost automorphic*. Functions are *almost automorphic* which are both right and left almost automorphic.

Remark. When convergence in (1.2.1) is required to be uniform, then sequences clearly suffice, and f is almost periodic in the sense of Bochner-von Neumann. ([15].) From this (1.2.2) follows. ([5].)

We shall deal at first with the set $A_L=A_L(G)$ of left almost automorphic functions on G. The reader will readily see that the mapping $t\to t^{-1}$ of G onto G interchanges the roles of left and right almost automorphy.

THEOREM 1.2.1. *The following properties are enjoyed by* A_L:

i) *If* $f_1, f_2\in A_L$, *then* f_1+f_2, $f_1\cdot f_2\in A_L$.

ii) *If* $f\in A_L$ *and* $\lambda\in \boldsymbol{C}$ (= *complex numbers*), *then* $\lambda f\in A_L$. *The constant function* λ *is in* A_L.

iii) *If* $f\in A_L$, *then* $\bar f\in A_L$ *where* $\bar f$ *is the complex conjugate of* f.

iv) *If* $f\in A_L$, *then* $|f|$ *is uniformly bounded on* G. ($|\cdot|$ *stands for absolute value.*)

v) *If* $\{f_n\}\subset A_L$ *is a sequence such that* $\lim f_n=f$ *holds uniformly on* G *for some function* f, *then* $f\in A_L$.

vi) *If* $f\in A_L$ *and* $\gamma\in G$, *then* $\gamma f, f\gamma\in A_L$.

Proof. Properties i)-iii) follow quite readily from the definitions. If property iv) did not hold, there would exist a sequence $\{\alpha_n\} \subset G$ such that $\lim_{n\to\infty} |f(\alpha_n)| = \infty$. For no subnet of $\{\alpha_n\}$ could (1.2.1) hold with $g(e)$ finite ($e =$ identity), and hence iv) must hold.

In v) note that f as the uniform limit of bounded functions is itself uniformly bounded. Therefore given a net $\alpha' = \{\alpha_i\}_{i\in\Lambda'}$ of group elements there will exist a subnet $\alpha = \{\alpha_i\}_{i\in\Lambda}$ such that $T_\alpha f$ and $T_{\alpha^{-1}}T_\alpha f$ exist. This last fact is standard and follows from the Tychonoff product theorem. We claim $T_{\alpha^{-1}}T_\alpha f = f$. To see this fix $\epsilon > 0$, and choose n so large that $\sup_{t\in G} |f_n(t) - f(t)| \leqq \epsilon$. Let a subnet β of α be so chosen that $T_\beta f_n$ exists and $T_{\beta^{-1}}T_\beta f_n = f_n$. Noting that $T_{\alpha^{-1}}T_\alpha = T_{\beta^{-1}}T_\beta f$ we have that

$$|T_{\alpha^{-1}}T_\alpha f - T_{\beta^{-1}}T_\beta f_n| = |T_{\beta^{-1}}T_\beta f - T_{\beta^{-1}}T_\beta f_n| = T_{\beta^{-1}}T_\beta |f - f_n| \leqq \epsilon.$$

By the triangle inequality

$$|T_{\alpha^{-1}}T_\alpha f - f| \leqq |T_{\alpha^{-1}}T_\alpha f - T_{\beta^{-1}}T_\beta f_n| + |T_{\beta^{-1}}T_\beta f_n - f| \leqq \epsilon + \epsilon = 2\epsilon.$$

Letting ϵ tend to zero gives the desired result

It is immediate for vi) that $f\gamma \in A_L$. For if $T_{\alpha^{-1}}T_\alpha f\cdot(s) = f(s)$ holds for $s \in G$, it holds for $s = t\gamma$, and so $T_{\alpha^{-1}}T_\alpha(f_\gamma) = f\gamma$. To see that $\gamma f \in A_L$ suppose given a net $\{\alpha_i'\}_{i\in\Lambda'}$ of group elements, and from this given net extract a subnet $\alpha = \{\alpha_i\}_{i\in\Lambda}$ such that both $\lim_{i\in\Lambda}(\gamma\alpha_i)f = g$ and $\lim(\alpha_i^{-1}\gamma^{-1})g = f$ hold. Then $g = \lim_{i\in\Lambda}\alpha_i(\gamma f)$, and we have $\lim_{i\in\Lambda}\alpha_i^{-1}g = \lim_{i\in\Lambda}(\alpha_i^{-1}\gamma^{-1}\gamma)g = \gamma f$. Thus $T_{\alpha^{-1}}T_\alpha\gamma f = \gamma f$, and $\gamma f \in A_L$. This completes the proof of our theorem.

Parts i)-v) may be restated as

THEOREM 1.2.1′. *The set A_L if given the norm $\|\cdot\|$, $\|f\| = \sup_{t\in G}|f(t)|$, is a commutative C^* algebra with unit.* ([13].)

It is not difficult using only property vi) of A_L to prove the existence of a normal subgroup G_0 of G such that each $f \in A_L$ is constant on the cosets of G_0, and such that if A_L is thought of as an algebra on G/G_0, then A_L separates the points of G/G_0. We shall not need this fact, although it will follow from results of Section 3.

1.3. Left and right almost automorphy. We will observe in this section that a left almost automorphic function is automatically right almost automorphic, a fact which is well known for the almost periodic functions. ([13].)

LEMMA 1.3.1. *If $f \in A_L$ and $T_\alpha f = g$ for some net α, then already $T_{\alpha^{-1}} g = f$.*

Proof. If $T_{\alpha^{-1}} g = f$ does not already hold, there must exist an $\epsilon > 0$ and $t \in G$ and a subnet $\beta = \{\beta_j\}_{j \in \Lambda}$ of α such that $|g(\beta_j^{-1} t) - f(t)| \geqq \epsilon$ for each $j \in \Lambda$. Since $T_\beta f = g$, it cannot be for any subnet γ of β that $T_{\gamma^{-1}} T_\gamma f = T_{\gamma^{-1}} T_\beta f = T_{\gamma^{-1}} g = f$. This contradicts the assumption that $f \in A_L$, and our lemma obtains.

THEOREM 1.3.1. *A function f on G is left almost automorphic if and only if it is right almost automorphic.*

Proof. Suppose $f \in A_L$. Then by Theorem 1.2.1 f is uniformly bounded on G. Given a net $\alpha' = \{\alpha_i'\}_{i \in \Lambda'}$ extract a subnet $\alpha = \{\alpha_i\}_{i \in \Lambda}$ such that $S_\alpha f = g$ and $S_{\alpha^{-1}} g = h$ exist for some functions g and h on G. We must prove $f = h$. To this end let $t \in G$ be a fixed element of G. Since $f \in A_L$, there is a subnet $t\beta^{-1} = \{t\beta_j^{-1}\}_{j \in \Lambda_1}$ of $\{t\alpha_j^{-1}\}_{j \in \Lambda}$ such that $T_{t\beta^{-1}} f = g_1$ and $T_{\beta t^{-1}} g_1 = f$ for some function g_1 on G. For each finite set $N \subset G$ of cardinality $|N|$ choose j, then k in Λ_1 such that both

$$\max_{s \in N} |f(t\beta_k^{-1}\beta_j t^{-1} s) - f(s)| < 1/|N| \tag{1.3.1}$$

and

$$|f(t\beta_j^{-1}\beta_k) - h(t)| < 1/|N|. \tag{1.3.2}$$

Let $\delta_N = t\beta_k^{-1}\beta_j t^{-1}$, $k = k(N)$, $j = j(N)$. Then $\{\delta_N\} = \delta$ is a net indexed by the system of finite subsets of G directed by inclusion. From (1.3.1) we have that $T_\delta f = f$. By Lemma 1.3.1 $T_{\delta^{-1}} f = f$. It follows from (1.3.2) that $T_{\delta^{-1}} f \cdot (t) = h(t)$, and so $f(t) = h(t)$. Since $t \in G$ is arbitrary $f \equiv h$, and the theorem is proved.

Theorem 1.3.1 allows us to deal with the algebra $A = A(G)$ of almost automorphic functions on G without loss of generality.

2. Bohr almost automorphy.

We introduce and discuss in this section the notion of a Bohr almost automorphic function. The principal result of the discussion, Theorem 2.2.1, asserts the equivalence of Bohr almost automorphy with Bochner almost automorphy.

2.1. Definitions and lemmas.

Definition 2.1.1. A subset E of a group G is *relatively dense* if there exist elements $s_1, \cdots, s_m$ and $t_1, \cdots, t_n$ in G such that $\bigcup_{i=1}^{m} s_i E = G = \bigcup_{j=1}^{n} E t_j$.

Definition 2.1.2. A *bounded* function f on a group G shall be called *Bohr almost automorphic* if for each finite set $N \subset G$ and prescribed $\epsilon > 0$ there is a set $B_\epsilon = B_\epsilon(N) \subset G$ such that

i) B_ϵ is relatively dense

ii) $B_\epsilon = B_\epsilon^{-1}$ where $B_\epsilon^{-1} = \{t^{-1} \mid t \in B_\epsilon\}$

iii) If $\tau \in B_\epsilon$, then $\max\limits_{s,t \in N} |f(s\tau t) - f(st)| < \epsilon$

iv) If $\tau_1, \tau_2 \in B_\epsilon$, then $\max\limits_{s,t \in N} |f(s\tau_1\tau_2^{-1}t) - f(st)| < 2\epsilon$.

We remark that if in iii) G is substituted for N, then iv) follows from a 2ϵ argument.

LEMMA 2.1.1. *Let f be Bochner almost automorphic on G. If $N \subset G$ is finite and $\epsilon > 0$ is given, the set*

$$C_\epsilon(N) = \{\tau \mid \max_{s,t \in N} |f(s\tau t) - f(st)| < \epsilon\} \tag{2.1.1}$$

is relatively dense.

Proof. Suppose for the sake of contradiction that there exists a finite set $N \subset G$ and an $\epsilon > 0$ such that the set $C_\epsilon(N)$ of (2.1.1) is not relatively dense. Then one (or both) of two cases occurs. Either a) there do not exist $s_1, \cdots, s_m \in G$ such that $\bigcup\limits_{i=1}^{m} s_i C_\epsilon = G$ or b) there do not exist $t_1, \cdots, t_n \in G$ such that $\bigcup\limits_{j=1}^{n} C_\epsilon t_j = G$. Considerations of symmetry show that it is sufficient to prove the latter to be impossible. We construct a sequence $\{t_n\} \subset G$ such that $t_n t_k^{-1} \notin C_\epsilon$ for $k < n$. For $n = 1$ let t_1 be an arbitrary element of G. Having chosen $t_1, \cdots, t_n \in G$ such that $t_m t_l^{-1} \notin C_\epsilon$, $m > l$ there will exist by the assumption that b) holds an element $t = t_{n+1}$ with $t_{n+1} \notin \bigcup\limits_{j=1}^{n} C_\epsilon t_j$. What is the same thing $t_{n+1} t_k^{-1} \notin C_\epsilon$ for $k \leqq n$. From the sequence so constructed extract a subnet $\{\alpha_i\}_{i \in \Lambda}$, $\alpha_i = t_{n(i)}$ such that $T_\alpha sft = g_{st}$ and $T_{\alpha^{-1}} g_{st} = sft$ for $s, t \in N$. There are but finitely many pairs $s, t \in N$ and $sft \in A$ by Theorem 1.2.1 so such a subnet surely exists. Let $\alpha_j = t_{n(j)}$ be chosen so that $\max\limits_{s,t \in N} |g_{st}(\alpha_j^{-1}) - f(st)| < \epsilon/2$, and then choose $\alpha_i = t_{n(i)}$ with $n(i) > n(j)$ such that

$$\max_{s,t \in N} |f(s\alpha_i\alpha_j^{-1}t) - f(st)| \leqq \max_{s,t \in N} |f(s\alpha_i\alpha_j^{-1}t) - g_{st}(\alpha_j^{-1})| + \max_{s,t \in N} |g_{st}(\alpha_j^{-1}) - f(st)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus $\alpha_i\alpha_j^{-1} = t_{n(i)} t_{n(j)}^{-1} \in C_\epsilon$ which since $n(i) > n(j)$ contradicts our con-

struction of the sequence $\{t_n\}$. Therefore C_ϵ must be relatively dense by the aformentioned consideration of symmetry and the preceding contradiction.

LEMMA 2.1.2. *Let f be Bochner almost automorphic on G. Given $\epsilon > 0$ and a finite set $N \subset G$ there exist $\delta > 0$ and a finite superset M of N such that whenever $\sigma, \tau \in C_\delta(M)$, then $\sigma^{-1}\tau \in C_\epsilon(N)$.*

Proof. Assume the contrary. Then for some finite set $N \subset G$ and $\epsilon > 0$ and every finite superset M of N and $\delta > 0$ there must exist $\sigma, \tau \in C_\delta(M)$ with $\sigma^{-1}\tau \notin C_\epsilon(N)$. Choose a sequence $\{\delta_n\}$ of positive real numbers decreasing to zero so fast that $\sum_{n=1}^{\infty} \delta_n < \infty$. Let a sequence $\{M_n\}$ of finite supersets of N together with a sequence (σ_n, τ_n) of pairs of elements be chosen as follows: Let $M_1' = N \cup \{e\}$, and set $M_1 = M_1' \cup (M_1')^{-1}$. M_1 is a finite superset of N, and so by assumption there exists $(\sigma, \tau) = (\sigma_1, \tau_1) \in G$ such that $\sigma, \tau \in C_{\delta_1}(M_1)$, but $\sigma^{-1}\tau \notin C_\epsilon(N)$. Having chosen $M_1, \cdots, M_k$ and $(\sigma_1, \tau_1) \cdots (\sigma_k, \tau_k)$ we set $M_{k+1}' = M_k \cdot M_k \cdot N_k$ where $N_k = \{e, \sigma_k, \tau_k, \sigma_k^{-1}\tau_k\}$, and define $M_{k+1} = M_{k+1}' \cup (M_{k+1}')^{-1}$. M_{k+1} is a finite superset of N, and again by assumption there exists $(\sigma, \tau) = (\sigma_{k+1}, \tau_{k+1})$ such that $\sigma_{k+1}, \tau_{k+1} \in C_{\delta_{k+1}}(M_{k+1})$, but $\sigma_{k+1}^{-1}\tau_{k+1} \notin C_\epsilon(N)$. The construction then proceeds by induction. Let $G_0 = \bigcup_{k=1}^{\infty} M_k$. If $x, y \in G_0$, then $x, y \in M_k$ for k sufficiently large implying by construction that $xy^{-1} \in M_{k+2} \subset G_0$. Thus G_0 is a subgroup of G, and it is an easy matter to verify that f when restricted to G_0 is again almost automorphic. A sequence $\{\alpha_k\}$ of group elements is now defined as follows: Let $\alpha_1 = \tau_1, \alpha_2 = \sigma_1$. For $k \geqq 1$ define $\alpha_{2k+1} = \prod_{n=1}^{k+1} \tau_n, \alpha_{2k+2} = \alpha_{2k-1}\sigma_{k+1}$. If $s, t \in M_k$ then $s\alpha_{2k-1} \in M_{k+1}$ since $s, \alpha_{2k-3} \in M_k$ and $\tau_k \in N_k$. We have

$$\begin{aligned} |f(s\alpha_{2k+1}t) - f(s\alpha_{2k+2}t)| &\leqq |f(s\alpha_{2k-1}\tau_{k+1}t) - f(s\alpha_{2k-1}t)| \\ &\quad + |f(s\alpha_{2k-1}t) - f(s\alpha_{2k-1}\sigma_{k+1}t) \leqq \delta_{k+1} + \delta_{k+1} = 2\delta_{k+1}. \end{aligned}$$

Therefore if we can show $\lim_{k\to\infty} f(s\alpha_{2k+1}t) = g(s,t)$ exists, then so will $\lim_{k\to\infty} f(s\alpha_k t) = g(s,t)$ hold. Now if $s, t \in M_k$ and $k < j$, $s\alpha_{2(k+i)+1} \in M_{k+i+2}$. Hence

$$\begin{aligned} |f(s\alpha_{2k+1}t) - f(s\alpha_{2j+1}t)| &\leqq \sum_{i=0}^{j-k-1} |f(s\alpha_{2(k+i)+1}t) - f(s\alpha_{2(k+i+1)+1}t)| \\ &= \sum_{i=0}^{j-k-1} |f(s\alpha_{2(k+i)+1}t) - f(s\alpha_{2(k+i)+1}\tau_{k+i+2}t)| \qquad (2.1.2) \\ &\leqq \sum_{i=0}^{j-k-1} \delta_{k+i+2}. \end{aligned}$$

The last sum on the right of (2.1.2) goes to zero independently of j as $k \to \infty$. Thus $\lim_{k\to\infty} f(s\alpha_{2k+1}t) = g(s,t)$, and by our previous remark $\lim_{k\to\infty} f(s\alpha_k t) = g(s,t)$. Since f is almost automorphic, lemma 1.3.1 implies that $\lim_{k\to\infty} g(s\alpha_k^{-1}, t) = f(st)$ holds for all $s, t \in G_0$. Let j be chosen so large that $\max_{s,t\in N} | g(s\alpha_{2j}^{-1}, t) - f(st) | < \epsilon/4$, and then choose $k > j$ so large that $\max_{s,t\in N} | f(s\alpha_{2j}^{-1}\alpha_{2k+1}t) - g(s\alpha_{2j}^{-1}, t) | < \epsilon/4$. Combining these last two inequalities

$$(2.1.3) \qquad \max_{s,t\in N} | f(s\alpha_{2j}^{-1}\alpha_{2k+1}t) - f(st) | < \epsilon/4 + \epsilon/4 = \epsilon/2$$

Now $\alpha_{2j}^{-1} = \sigma_{j+1}^{-1}\tau_j^{-1}\tau_{j-1}^{-1} \cdots \tau_1^{-1}$, and $\alpha_{2k+1} = \tau_1 \cdots \tau_{k+1}$ so since $k > j$, $\alpha_{2j}^{-1}\alpha_{2k+1} = \sigma_{j+1}^{-1}\tau_{j+1} \cdots \tau_{k+1}$. If $s, t \in N$, then

$$s_{j+1}^{-1}\tau_{j+1} \in M_{j+2}, s\sigma_{j+1}^{-1}\tau_{j+1}\tau_{j+2}. \in M_{j+3}, \cdots, s\sigma_{j+1}^{-1}\tau_{j+1} \cdots \tau_k \in M_{k+1}.$$

Hence we have

$$\begin{aligned} &\max_{s,t\in N} | f(s\alpha_{2j}^{-1}\alpha_{2k+1}t) - f(s\sigma_{j+1}^{-1}\tau_{j+1}t) | \\ (2.1.4) \qquad &\leqq \max_{s,t\in N} \sum_{i=0}^{k-j-1} | f(s\sigma_{j+1}^{-1}\tau_{j+1} \cdots \tau_{j+1+i}t) - f(s\sigma_{j+1}^{-1}\tau_{j+1} \cdots \tau_{j+2+i}t) | \\ &\leqq \sum_{i=0}^{k-j-1} \delta_{j+i+1}. \end{aligned}$$

Let j be chosen so large that $\sum_{n=j}^{\infty} \delta_n < \epsilon/2$ which is possible since $\sum_{n=0}^{\infty} \delta_n < \infty$. It then follows from (2.1.3), (2.1.4), and the triangle inequality that $\max_{s,t\in N} | f(st) - f(s\sigma_{j+1}^{-1}\tau_{j+1}t) | < \epsilon$, and this inequality contradicts the fact that $(\sigma_{j+1}, \tau_{j+1})$ were so chosen that $\sigma_{j+1}^{-1}\tau_{j+1} \notin C_\epsilon(N)$. This completes the proof of the lemma.

For later applications it will be useful to have the following strengthening of Lemma 2.1.2.

COROLLARY 2.1.2. *Let f, N, $\epsilon > 0$ be as in Lemma 2.1.2. For any integer $n > 0$ there exists a finite superset M of N and a $\delta > 0$ such that if $\tau_1, \cdots, \tau_n \in C_\delta(M)$, then $\tau_1^{\epsilon_1} \cdots \tau_n^{\epsilon_n} \in C_\epsilon(N)$ for any choices of $\epsilon_i = 0, 1$,* or -1.

Proof. The proof is by induction on n. For $n = 1$ select a set $M \supset N$ and $\delta > 0$ by the previous lemma. If $\tau \in C_\delta(M)$, then $\tau^{\epsilon_1} \in C_\epsilon(N)$ for $\epsilon_1 = 0$, 1, or -1 since $e \in C_\delta(M)$, and $\tau^{-1}e \in C_\epsilon(N)$. Suppose the corollary holds for some integer n, and let $M_1 \supset N$ and $\delta_1 > 0$ be chosen by Lemma 2.1.2 so that whenever $\tau, \sigma \in C_{\delta_1}(M_1)$, then $\sigma^{-1}\tau \in C_\epsilon(N)$. Using Lemma

2.1.2 once more let $M_2 \supset M_1$ and $\delta_2 > 0$ be choosen so that when $\sigma, \tau \in C_{\delta_2}(M_2)$, then $\sigma^{-1}\tau \in C_{\delta_1}(M_1)$. Note that if $\tau \in C_{\delta_2}(M_2)$, then $\sigma^{-1} \in C_{\delta_1}(M_1)$. By our induction assumption we choose a set $M \supset M_2$ and $\delta > 0$ such that whenever $\tau_1, \cdots, \tau_n \in C_\delta(M)$, then $\tau_1^{\epsilon_1} \cdots \tau_n^{\epsilon_n} \in C_{\delta_2}(M_2)$ for any choice of $\epsilon_j = 0, 1$, or -1. Let $\tau_1, \cdots, \tau_{n+1}$ be elements of $C_\delta(M)$, and suppose $\gamma = \tau_1^{\epsilon_1} \cdots \tau_{n+1}^{\epsilon_{n+1}} = \gamma'\gamma''$ where $\gamma' = \tau_1^{\epsilon_1} \cdots \tau_n^{\epsilon_n}$, $\gamma'' = \tau_{n+1}^{\epsilon_{n+1}}$ with again $\epsilon_j = 0, 1$, or -1. Then both $\gamma' \in C_{\delta_2}(M_2)$ and $\gamma'' \in C_{\delta_2}(M_2)$. By our choice of M_2, δ_2

$$(\gamma')^{-1} \in C_{\delta_1}(M_1), \qquad \gamma'' \in C_{\delta_1}(M_1),$$

and finally $\gamma = ((\gamma')^{-1})^{-1}\gamma'' \in C_\epsilon(N)$. The corollary then follows by induction.

2.2. Equivalence of Bochner and Bohr almost automorphy.

THEOREM 2.2.1. *A function f on G is Bochner almost automorphic if and only if it is Bohr almost automorphic.*

Proof. Let f be Bohr almost automorphic and $\alpha' = \{\alpha_i'\}_{i \in \Lambda'}$ be a net of group elements. By definition f is bounded on G, and so there is a subnet $\{\alpha_i\}_{i \in \Lambda} = \alpha$ or α' such that $T_\alpha f = g$ and $T_{\alpha^{-1}} g = h$ exist. Fixing $t \in G$ and $\epsilon > 0$ choose a set $B_\epsilon = B_\epsilon\{e, t\}$ satisfying conditions i)-iv) of Definition 2.1.2. Since B_ϵ is relatively dense, there exist elements $s_1, \cdots, s_m$ of G such that each $s \in G$ may be written $s = \tau s_j$ where $\tau \in B_\epsilon$ and $1 \leqq j \leqq m$. For each $i \in \Lambda$ we can write $\alpha_i = \tau_i s_j$, $j = j(i)$. There are but finitely many s_j, so there will exist a subnet $\beta = \{\beta_k\}_{k \in \Lambda_1}$, $\beta_k = \tau_k s_{j_0}$ where j_0 is independent of k. It remains true for β that $T_{\beta^{-1}} T_\beta f = h$. Let k then i be chosen from Λ_1 so large that $|f(\beta_i \beta_k^{-1} t) - h(t)| < \epsilon$. Observing that $\beta_i \beta_k^{-1} = \tau_i \tau_k^{-1}$ and using iv) of Definition 2.1.2 we have $|f(\beta_i \beta_k^{-1} t) - f(t)| < 2\epsilon$. Thus by the triangle inequality $|h(t) - f(t)| < 3\epsilon$. Since ϵ then t are arbitrary $f = h$ and f is Bochner almost automorphic as was to be proved.

Suppose conversely that f is Bochner almost automorphic. First f is bounded by Theorem 1.2.1. Given a finite set $N \subset G$ and $\epsilon > 0$ choose a finite superset M of N and $\delta > 0$ for $n = 2$ in Corollary 2.1.2. Define $B_\epsilon(N) = C_\delta(M) \cup C_\delta(M)^{-1}$. Then B_ϵ is relatively dense by Lemma 2.1.1 and $B_\epsilon = B_\epsilon^{-1}$ by definition. If $\tau \in B_\epsilon$ then $\tau \in C_\epsilon(N)$ while if $\sigma, \tau \in B_\epsilon$ then $\sigma^{-1}\tau \in C_\epsilon(N)$ using Corollary 2.1.2. Thus f enjoys properties i)-iv) of Definition 2.1.2 in addition to being bounded, and this completes the proof that f is Bohr almost automorphic.

2.3. Almost periodic functions on the integers. To illustrate the usefulness of Theorem 2.2.1. we will give here a direct proof for the group

$G = \boldsymbol{Z}$ of integers of a theorem which will be seen later to hold for an arbitrary group.

THEOREM 2.3.1. *A function f on $\boldsymbol{Z}$ is almost periodic if and only if it is almost automorphic and $T_\alpha f$ is almost automorphic when the limit exists.*

Proof. The "only if" part follows from the fact that $T_\alpha f$ when it exists is almost periodic and in particular almost automorphic. Suppose that f is almost automorphic but not almost periodic. We will prove the existence of a sequence $\{\alpha_k\} = \alpha$ of integers such that $T_\alpha f$ exists but is not almost automorphic. To this end let $\{\epsilon_k\}$ be a sequence of positive real numbers decreasing to zero, and for each $k = 1, 2, \cdots$ let $I_k = \{n \mid -k \leqq n \leqq k\} \subset \boldsymbol{Z}$. By Theorem 2.2.1 there exists for each k a set $B_{\epsilon_k} = B_{\epsilon_k}(I_k)$ satisfying i)-iv) of Definition 2.1.2 with $N = I_k$ and $\epsilon = \epsilon_k$. Because f is not almost periodic there must exist an $\epsilon > 0$ such that $C_\epsilon(\boldsymbol{Z}) = \{\tau \mid |f(s+\tau) - f(s)| < \epsilon, s \in \boldsymbol{Z}\}$ is not relatively dense. We assume $\epsilon > \epsilon_k$, $k = 1, 2, \cdots$ by discarding finitely many k if necessary. There must exist sequences $\{s_k\}$ and $\{t_k\}$ of integers with $s_k \in B_{\epsilon_k}$ such that $|f(s_k + t_k) - f(t_k)| \geqq \epsilon$ for each k. Using the familiar Cantor diagonal procedure choose a sequence $\{k_n\}$ of indices such that both $\lim_{n\to\infty} f(x + s_{k_n} + t_{k_n}) = g(x)$ and $\lim_{n\to\infty} f(x + t_{k_n}) = h(x)$ exist for each $x \in \boldsymbol{Z}$. Then $|g(0) - h(0)| \geqq \epsilon$ and consequently $g \neq h$. For k fixed write $t_{k_n} = \tau_{n,k} + r_{n,k}$ with $\tau_{n,k} \in B_{\epsilon_k}$ and $|r_{n,k}| \leqq R_k < \infty$ using the fact that B_{ϵ_k} is relatively dense. By a further refinement if necessary assume that $r_{n,k} = r_k$ is independent of n. (We will not renumber the sequences.) If $k_n \geqq k$ and $|x + r_k| \leqq k$, it must be that

$$\begin{aligned}
|f(x + s_{k_n} + t_{k_n}) - f(x + t_{k_n})| &\leqq |f(x + s_{k_n} + t_{k_n}) - f(x + r_k)| \\
&\quad + |f(x + r_k) - f(x + t_{k_n})| = |f(x + r_k + \tau_{n,k} + s_{k_n}) - f(x + r_k)| \\
&\quad + |f(x + r_k) - f(x + r_k + \tau_{k,n})| \leqq 2\epsilon_k + \epsilon_k = 3\epsilon_k.
\end{aligned}$$

(Since $-s_{k_n} \in B_{\epsilon_k}$, then $\tau_{n,k} + s_{k_n} \in C_{\epsilon_k}(I_k)$.) Letting n tend to infinity we have $|g(x) - h(x)| \leqq 3\epsilon_k$ when $x + r_k \leqq k$. Let a subsequence $\{r_{k_m}\} \subset \{r_k\}$ be chosen so that both $\lim_{m\to\infty} g(x - r_{k_m}) = p(x)$ and $\lim_{m\to\infty} f(x - r_{k_m}) = q(x)$ exist for all $x \in \boldsymbol{Z}$. If $|x| \leqq k_m$, then $|x - r_{k_m} + r_{k_m}| \leqq k_m$, and it follows that $|g(x - r_{k_m}) - h(x - r_{k_m})| \leqq 3\epsilon_{k_m}$. Letting $m \to \infty$ it follows that $p(x) \equiv q(x)$. Since $g \neq h$ it cannot be that both $\lim_{m\to\infty} p(x + r_{k_m}) = g(x)$ and $\lim_{m\to\infty} q(x + r_{k_m}) = h(x)$, and hence one or both of g, h is not almost automorphic. The proof is complete.

3. Almost automorphic flows.

3. 1. The flows associated with a function. We fix for the discussion an almost automorphic function f on the group G with $\sup_{t\in G}|f(t)| = M < \infty$. If $I = \{z \in \boldsymbol{C} \mid |z| \leqq M\}$, let $\tilde{X} = I^{G\times G} = \prod_{(s,t)\in G\times G} I_{(s,t)}$ where $I_{(s,t)} = I$ for each $(s,t) \in G \times G$. Corresponding to f in the compact space $\tilde{X}$ will be the point ξ whose (s,t)-th coordinate $\xi(s,t)$ is $f(st)$. Let X' be the set of points η_γ in X of the form $\eta_\gamma(s,t) = \xi(s\gamma,t) = \xi(s,\gamma t)$. For each $\gamma \in G$ operations L_γ, R_γ are defined on X' coordinate-wise by $L_\gamma\eta\cdot(s,t) = \eta(s\gamma,t)$ and $R_\gamma\eta\cdot(s,t) = \eta(s,\gamma t)$. Thus $X' = \bigcup_{\gamma\in G} L_\gamma\xi = \bigcup_{\gamma\in G} R_\gamma\xi$. Note that $L_{\gamma_1}L_{\gamma_2} = L_{\gamma_1\gamma_2}$ and $R_{\gamma_1}R_{\gamma_2} = R_{\gamma_2\gamma_1}$. L_γ and R_γ extend to be continuous on the closure X of X', and G is there represented in two ways by homeomorphisms of X. The respective representations are called the flows (G_L, X) and (G_R, X) arising from the function f.

If $\alpha = \{\alpha_i\}_{i\in\Lambda}$ is a net of group elements such that $\lim_{i\in\Lambda} L_{\alpha_i}x = y$ exists, we shall write $y = T_\alpha x$. If $\lim_{i\in\Lambda} R_{\alpha_i}x = y$ holds, then $y = S_\alpha x$.

Definition 3.1.1. A point $x \in X$ is *left almost automorphic* if $T_{\alpha^{-1}}T_\alpha x = x$ whenever the limits exist. If $S_{\alpha^{-1}}S_\alpha x = x$ whenever the limits exist, then x is *right almost automorphic.* Points which are both right and left almost automorphic are *almost automorphic.*

Since f and all of its translates γf, $f\gamma$ are almost automorphic, the points ξ, $L_\gamma\xi$, $R_\gamma\xi$ are also almost automorphic.

LEMMA 3.1.1. *If $x \in X$ and $\alpha = \{\alpha_i\}_{i\in\Lambda}$ are such that $T_{\alpha^{-1}}x$ and $S_{\alpha^{-1}}x$ exist, and if $T_{\alpha^{-1}}x = \xi$, then $S_{\alpha^{-1}}x = \xi$. (By symmetry if $S_{\alpha^{-1}}x = \xi$, then $T_{\alpha^{-1}}x = \xi$.)*

Proof. By definition of X as the closure in $\tilde{X}$ of X' there will exist a net $\gamma = \{\gamma_j\}_{j\in\Lambda_1}$ such that $T_\gamma\xi = x$. Given that $T_{\alpha^{-1}}T_\gamma\xi = \xi$ and $S_{\alpha^{-1}}T_\gamma\xi = \eta$ for some $\eta \in X$, it is to be shown that $\xi = \eta$. Given a finite set $N \subset G$ and $\epsilon > 0$ let a finite superset M of N and $\delta > 0$ be chosen by Corollary 2.1.2 for $n = 3$. By Lemma 2.1.1 $C_\delta(M)$ is relatively dense, and so for each $i \in \Lambda_1$ γ_i may be written $\gamma_i = \tau_i r_i$ where $\tau_i \in C_\delta(M)$ and $r_i \in R$ for some finite set R. By choosing a subnet $\gamma_1 = \{\gamma_{i(j)}\}_{j\in\Lambda_2}$ if necessary we may assume $\gamma_{i(j)} = \tau_{i(j)}r$ where r is independent of j. Since $T_{\alpha^{-1}}T_{\gamma_1}\xi = \xi$, choose $k \in \Lambda$ and $j \in \Lambda_2$ such that $r\alpha_k^{-1}\gamma_{i(j)}r^{-1} = \tau_{kj} \in C_\delta(M)$. Solving for α_k^{-1}, $\alpha_k^{-1} = r^{-1}\tau_{kj}r\gamma_{i(j)}^{-1}$ and $\gamma_{i(j)}\alpha_k^{-1} = \gamma_{i(j)}r^{-1}\tau_{kj}r\gamma_{i(j)}^{-1} = \tau_{i(j)}\tau_{kj}\tau_{i(j)}^{-1}$. Since $\tau_{i(j)}, \tau_{kj} \in C_\delta(M)$, it follows

from Corollary 2.1.2 (our choice of M, δ) that $\max_{s,t \in N} |f(s\gamma_{i(j)}\alpha_k^{-1}t) - f(st)| < \epsilon$. That is $\max_{s,t \in N} |R_{\alpha_k^{-1}} L_{\gamma_{i(j)}} \xi \cdot (s,t) - \xi(s,t)| < \epsilon$. Passing to the limit in j then k gives us that $\max_{s,t \in N} |\eta(s,t) - \xi(s,t)| \leqq \epsilon$. Since ϵ, then N were arbitrary $\xi = \eta$ and the lemma is proved.

Let us remark here that if $T_{\alpha^{-1}} x = \xi$ holds, then for any subnet β of α for which $S_{\beta^{-1}} x$ exists, it must be that $\xi = S_{\beta^{-1}} x$ by the preceding lemma. Therefore the only cluster point (and it must have one) of the net $\{R_{\alpha_j^{-1}} x\}_{j \in \Lambda}$ is ξ and already $S_{\alpha^{-1}} x = \xi$.

LEMMA 3.1.2. *Let $x, y \in X$ be such that there exists a net $\{\alpha_i\}_{i \in \Lambda}$ with $T_{\alpha^{-1}} T_\alpha x = y$. If $x = T_\beta \xi$ for a net $\beta = \{\beta_j\}_{j \in \Lambda_1}$, then $S_{\beta^{-1}} y = \xi$.*

Proof. We must show that $S_{\beta^{-1}} T_{\alpha^{-1}} T_\alpha T_\beta \xi = \xi$, and for this it is sufficient to show as we have earlier remarked that for any subnet $\gamma = \{\gamma_k\}_{k \in \Lambda_2}$ of β such that $S_{\gamma^{-1}} T_{\alpha^{-1}} T_\alpha T_\beta \xi = \xi_1$ exists, $\xi = \xi_1$. Note that $S_{\gamma^{-1}} T_{\alpha^{-1}} T_\alpha T_\beta \xi = S_{\gamma^{-1}} T_{\alpha^{-1}} T_\alpha T_\gamma \xi$. Fixing a finite set $N \subset G$ and $\epsilon > 0$ choose a finite set M containing N and a $\delta > 0$ for $n = 4$ in Corollary 2.1.2. For each $i \in \Lambda$ write $\alpha_i = r_i \tau_i$, $\tau_i \in C_\delta(M)$ and $r_i \in R$ where R is finite. A subnet $\{\alpha_i^1\}_{i \in \Lambda_3}$ of α is chosen so that $\alpha_i^1 = r\tau_i$ where $\tau_i \in C_\delta(M)$ and r is independent of $i \in \Lambda_3$. By a similar argument choose a subnet γ^1 of γ such that if $\gamma^1 = \{\gamma_k^1\}_{k \in \Lambda_4}$, then $\gamma_k^1 = \tau_k s$, $\tau_k \in C_\delta(M)$ and s independent of $k \in \Lambda_4$. Choose i, then j, then k, then l in Λ_4, Λ_3, Λ_3, Λ_4 respectively such that

$$\max_{s,t \in N} |\xi(s\alpha_j^{-1}\alpha_k\gamma_l, \gamma_i^{-1}t) - \xi_1(s,t)| < \epsilon.$$

By definition of $\xi(s,t)$ as $f(st)$ and the fact that $\alpha_j^{-1}\alpha_k = \tau_j^{-1}\tau_k$ and $\gamma_l\gamma_i^{-1} = \tau_l\tau_i^{-1}$ we have $\max_{s,t \in N} |\xi(s\alpha_j^{-1}\alpha_k\gamma_l, \gamma_i^{-1}t) - \xi(s,t)| < \epsilon$. By the triangle inequality $\max_{s,t \in N} |\xi(s,t) - \xi_1(s,t)| < 2\epsilon$, and since ϵ then N were arbitrary, $\xi = \xi_1$ as was to be proved.

COROLLARY 3.1.2. *Under the hypotheses of Lemma 3.1.2 we have $T_{\beta^{-1}} y = \xi$.*

Proof. Since $S_{\beta^{-1}} y = \xi$ we have from Lemma 3.1.1 that $T_{\beta^{-1}} y = \xi$.

LEMMA 3.1.3. *Let $x, y, z \in X$ be given by $T_\alpha \xi$, $T_\beta \xi$, $T_\gamma \xi$ respectively for certain nets $\alpha = \{\alpha_i\}_{i \in \Lambda_1}$, $\beta = \{\beta_j\}_{j \in \Lambda_2}$, $\gamma = \{\gamma_k\}_{k \in \Lambda_3}$. If $T_{\beta^{-1}} T_\alpha \xi = \xi$ and $T_{\beta^{-1}} T_\gamma \xi = \xi$, then also $T_{\alpha^{-1}} T_\gamma \xi = \xi$.*

Proof. It will suffice as in the argument of the preceding lemma to demonstrate that if $\delta = \{\delta_l\}_{l \in \Lambda_4}$ is a subnet of α such that $T_{\delta^{-1}} T_\gamma \xi = \xi_1$ exists,

then $\xi_1 = \xi$. Let N be a finite set and $\epsilon > 0$ a positive number. By Corollary 2.1.2 for $n = 2$ we choose a finite superset M of N and $\epsilon_0 > 0$ such that whenever $\tau_1, \tau_2 \in C_{\epsilon_0}(M)$, then $\tau_1\tau_2^{-1} \in C_\epsilon(N)$. Select first $i \in \Lambda_2$, then $j \in \Lambda_4$, then $k \in \Lambda_8$ so that

1) $$\max_{s,t \in M} | \xi(s\beta_i^{-1}\delta_j, t) - \xi(s,t) | < \epsilon_0,$$

2) $$\max_{s,t \in M} | \xi(s\beta_i^{-1}\gamma_k, t) - \xi(s,t) | < \epsilon_0,$$

3) $$\max_{s,t \in N} | \xi(s\delta_j^{-1}\gamma_k, t) - \xi_1(s,t) | < \epsilon.$$

By our choice of M, ϵ_0 it is also true from 1) and 2) that

$$\max_{s,t \in N} | \xi(s\delta_j^{-1}\gamma_k, t) - \xi(s,t) | < \epsilon.$$

Applying the triangle inequality to 3) and this last inequality we get $| \xi_1(s,t) - \xi(s,t) | < 2\epsilon$, $s, t \in N$. Letting ϵ tend to zero and N vary we see that $\xi = \xi_1$ as was to be proved.

Definition 3.1.2. We say $x \sim y$ for $x, y \in X$ if there exist nets α, β such that $T_{\alpha^{-1}}T_\alpha x = y$ and $S_{\beta^{-1}}S_\beta x = y$.

THEOREM 3.1.1. *The relation $x \sim y$ is an equivalence relation. Furthermore if there exists a net α such that $T_{\alpha^{-1}}T_\alpha x = y$, then there must exist a net β with $S_{\beta^{-1}}S_\beta x = y$, i.e., $x \sim y$. (By symmetry if there exists a net β with $S_{\beta^{-1}}S_\beta x = y$, then $x \sim y$.)*

Proof. Obviously $x \sim x$ so $\sim$ is reflexive. If $x \sim y$, then $T_{\alpha^{-1}}T_\alpha x = y$, $S_{\beta^{-1}}S_\beta x = y$ for certain nets α and β. For some net γ, $x = T_\gamma\xi$, and Corollary 3.1.2 implies that $T_{\gamma^{-1}}y = \xi$. Then $T_\gamma T_{\gamma^{-1}}y = x$. By Lemma 3.1.1, $S_{\gamma^{-1}}y = \xi$, and therefore $S_\gamma S_{\gamma^{-1}}y = x$. Thus $y \sim x$, and $x \sim y$ is symmetric. Suppose now that $x \sim y$ and $y \sim z$ where $x = T_\alpha\xi$, $y = T_\beta\xi$, and $z = T_\gamma\xi$. By Corollary 3.1.2, $\xi = T_{\beta^{-1}}x = T_{\beta^{-1}}z$, and by Lemma 3.1.3, $\xi = T_{\alpha^{-1}}z$. From Lemma 3.1.1, it again follows that $S_{\alpha^{-1}}z = \xi$. Finally $S_\alpha S_{\alpha^{-1}}z = x = T_\alpha T_{\alpha^{-1}}z$ and $z \sim x$. By symmetry $x \sim z$ and the relation $\sim$ is transitive. For the last statement of the theorem suppose $T_{\alpha^{-1}}T_\alpha x = y$ and $x = T_\gamma\xi$. Then $\xi = T_{\gamma^{-1}}y$ by Corollary 3.1.2 and $\xi = S_{\gamma^{-1}}y$ by Lemma 3.1.1. Then $T_\gamma T_{\gamma^{-1}}y = x = S_\gamma S_{\gamma^{-1}}y$ and $y \sim x$. By symmetry $x \sim y$ and the theorem is proved.

LEMMA 3.1.4. *Given a finite set $N \subset G$ and an $\epsilon > 0$ there is a finite superset M of N and a $\delta > 0$ such that whenever $\max_{s,t \in N} | x(s,t) - \xi(s,t) | < \delta$ and $x \sim y$, then $\max_{s,t \in N} | y(s,t) - \xi(s,t) | \leqq \epsilon$.*

Proof. Let $M \supset N$ and $\delta > 0$ be chosen by Corollary 2.1.2 for $n = 2$. Suppose $x \in X$ is such that $|x(s,t) - \xi(s,t)| < \delta$ for $s, t \in M$, and let $y \sim x$ be given by $y = T_{\beta^{-1}}T_{\beta}x$, $\beta = \{\beta_i\}_{i \in \Lambda_1}$. We assume there has been a preliminary refinement of β so that also $T_{\beta^{-1}}T_{\beta}\xi = \xi$. There will exist a net $\gamma = \{\gamma_j\}_{j \in \Lambda_2}$ such that $x = T_{\gamma}\xi$, and for some $j' \in \Lambda_2$ it will be true that if $k \geqq j'$, then $\max\limits_{s,t \in N} |\xi(s\gamma_k, t) - \xi(s,t)| < \delta$, the coordinate functions being continuous. In other words for $k \geqq j'$ $\gamma_k \in C_{\delta}(M)$. Let i then j in Λ_1 and then $k \geqq j'$ in Λ_2 be chosen so large that

a) $\max\limits_{s,t \in N} |\xi(s\beta_i^{-1}\beta_j\gamma_k, t) - y(s,t)| < \epsilon_0$ where $\epsilon_0 > 0$ is arbitrary,

b) $\max\limits_{s,t \in M} |\xi(s\beta_i^{-1}\beta_j, t) - \xi(s,t)| < \delta$. Then since $\beta_i^{-1}\beta_j \in C_{\delta}(M)$ and

$\gamma_k \in C_{\delta}(M)$ it follows from our choice of M that

c) $\max\limits_{s,t \in N} |\xi(s\beta_i^{-1}\beta_j\gamma_k, t) - \xi(s,t)| < \epsilon$.

Applying the triangle inequality to a) and c) we obtain

$$\max_{s,t \in N} |\xi(s,t) - y(s,t)| < \epsilon + \epsilon_0.$$

Since ϵ_0 can be made arbitrarily small,

$$\max_{s,t \in N} |\xi(s,t) - y(s,t)| \leqq \epsilon$$

as was to be proved.

Let us remark before proceeding that if $x \sim y$, then for each $\gamma \in G$, $L_{\gamma}x \sim L_{\gamma}y$ and $R_{\gamma}x \sim R_{\gamma}y$. For if $S_{\beta^{-1}}S_{\beta}x = y$, then $S_{\beta^{-1}}S_{\beta}L_{\gamma}x = L_{\gamma}y$, and by Theorem 3.1.1, $L_{\gamma}x \sim L_{\gamma}y$. Similarly if $T_{\alpha^{-1}}T_{\alpha}x = y$, then $T_{\alpha^{-1}}T_{\alpha}R_{\gamma}x = R_{\gamma}y$, and again by Theorem 3.1.1, $R_{\gamma}x \sim R_{\gamma}y$.

LEMMA 3.1.5. *Let $\{x_i\}_{i \in \Lambda}$ and $\{y_i\}_{i \in \Lambda}$ be nets such that* a) $x_i \sim y_i$ *for each $i \in \Lambda$ and* b) $\lim\limits_{i \in \Lambda} x_i = x$ *and* $\lim\limits_{i \in \Lambda} y_i = y$ *exist. Then $x \sim y$.*

Proof. For some net $\alpha = \{\alpha_i\}_{i \in \Lambda}$ $T_{\alpha}\xi = x$, $T_{\alpha^{-1}}x = \xi$. Fixing a finite set N and $\epsilon > 0$ choose a finite set $M \supset N$ and a $\delta > 0$ by the preceding lemma in such a way that if

$$\max_{s,t \in M} |x_1(s,t) - \xi(s,t)| < \delta$$

and $x_2 \sim x_1$, then

$$\max_{s,t \in N} |x_2(s,t) - \xi(s,t)| < \epsilon.$$

Let $j' \in \Lambda_1$ be such that for $j > j'$

$$|x(s\alpha_j^{-1}, t) - \xi(s,t)| < \delta,$$

$s, t \in M$. Since $x_i \to x$, there will be an index $i' \in \Lambda$ such that for $i > i'$ $\max_{s,t \in M} |x_i(s\alpha_j^{-1}, t) - \xi(s,t)| < \delta$ by the continuity of $L_{\alpha_j^{-1}}$. By the assumption that $x_i \sim y_i$, Lemma 3.1.4, and the remark following Lemma 3.1.4 we have $\max_{s,t \in N} |y_i(s\alpha_j^{-1}, t) - \xi(s,t)| < \epsilon$ for $i > i'$, $j > j'$. Passing to the limit in i gives us that $\max_{s,t \in N} |y(s\alpha_j^{-1}, t) - \xi(s,t)| \leqq \epsilon$ for $j > j'$. Since ϵ then N are arbitrary it follows that $T_{\alpha^{-1}}y = \xi$ and $T_\alpha T_{\alpha^{-1}}y = x$. By Theorem 3.1.1, $x \sim y$, and the lemma is proved.

As a corollary to Lemma 3.1.5 we can state

LEMMA 3.1.6. *Let U be a closed subset of X. Define $\tilde{U} = \bigcup_{x \in U} [x]$ where $[x] = \{y \in X \mid y \sim x\}$. Then $\tilde{U}$ is closed.*

Proof. Suppose $\{x_i'\}_{i \in \Lambda'}$ is a net in $\tilde{U}$ such that $\lim_{i \in \Lambda'} x_i' = x$ exists. For each $i \in \Lambda'$ there is by definition of $\tilde{U}$ a point $y_i' \in U$ with $y_i' \sim x_i'$. Since X is compact we may select a subnet $\{y_i\}_{i \in \Lambda}$ of $\{y_i'\}_{i \in \Lambda'}$ such that $\lim_{i \in \Lambda} y_i = y$ exists. Then $y \in U$ since U is closed, and since $y_i \sim x_i$ for each i, Lemma 3.1.5 implies that $y \sim x$. Thus $x \in \tilde{U}$, and $\tilde{U}$ is closed.

3.2. The group X_0. We let X_0 be the space $X/\sim$ of equivalence classes for $\sim$ with projection $\pi\colon X \to X_0$ given by $\pi x = [x]$. X_0 is given the strongest topology which is consistent with the continuity of π. A set $\tilde{U} \subset X_0$ will therefore be closed if and only if $\pi^{-1}\tilde{U} = U$ is closed in X. Lemma 3.1.6 implies that if $U \subset X$ is closed, then $\pi^{-1}\pi U$ is closed in X, and therefore $\pi U = \tilde{U}$ is closed in X_0. Hence π is a closed map and it follows (cf. [11]) that X_0 is a (compact) Hausdorff space. The remainder of the present section will be devoted to demonstrating that X_0 in a natural way carries the structure of a compact group. In the light of Theorem 2.3.1 this is not surprising since by forming X_0 we have by definition rid ourselves of all non-almost automorphic points.

Before proceeding we remark that if α, β, γ are nets such that a) $T_\alpha\xi$, $T_\beta\xi$ exist and $T_\alpha\xi \sim T_\beta\xi$ and b) $T_\gamma T_\alpha\xi$, $T_\gamma T_\beta\xi$ exist, then $T_\gamma T_\alpha\xi \sim T_\gamma T_\beta\xi$. For if $\gamma = \{\gamma_i\}_{i \in \Lambda}$, then for each $i \in \Lambda$, $L_{\gamma_i} T_\alpha\xi \sim L_{\gamma_i} T_\beta\xi$ and Lemma 3.1.5 implies that $T_\gamma T_\alpha\xi \sim T_\gamma T_\beta\xi$.

LEMMA 3.2.1. *Let $x, y, z \in X$ be given by $T_\alpha\xi$, $T_\beta\xi$, and $T_\gamma\xi$ respectively*

for nets $\alpha = \{\alpha_i\}_{i \in \Lambda_1}$, $\beta = \{\beta_j\}_{j \in \Lambda_2}$, *and* $\gamma = \{\gamma_k\}_{k \in \Lambda_3}$. *If* $x \sim y$ *and both* $T_\alpha z$, $T_\beta z$ *exist, then* $T_\alpha z \sim T_\beta z$.

Proof. We assume α, β, and γ have undergone a preliminary refinement so that $S_{\gamma^{-1}} T_{\alpha^{-1}} T_\beta T_\gamma \xi = \xi_1$ exists. Given a finite set $N \subset G$ and an $\epsilon > 0$ choose for $n = 2$ in Corollary 2.1.2 a finite superset M of N and $\delta > 0$ fulfilling the statement of that corollary. By our usual relative density argument there will exist subnets of α, β, and γ (which we will not reletter) such that we may choose l then i and j then k in Λ_3, Λ_1, Λ_2, Λ_3 respectively so that

1) $\max_{s,t \in N} |\xi(s\alpha_i^{-1}\beta_j\gamma_k, \gamma_l^{-1}t) - \xi_1(s,t)| < \epsilon$ and

2) $\alpha_i^{-1}\beta_j \in C_\delta(M)$ and $\gamma_k\gamma_l^{-1} \in C_\delta(M)$. (The first part of 2) follows from the fact that $T_{\alpha^{-1}} T_\beta \xi = \xi$.) Using our assumption on M and δ we obtain $\max_{s,t \in N} |\xi(s\alpha_i^{-1}\beta_j\gamma_k, \gamma_l^{-1}t) - \xi(s,t)| < \epsilon$. Thus $\max_{s,t \in N} |\xi(s,t) - \xi_1(s,t)| < 2\epsilon$, and letting ϵ tend to zero and N vary it follows that $\xi = \xi_1$. We have finally that $S_\gamma S_{\gamma^{-1}}(T_{\alpha^{-1}} T_\beta T_\gamma \xi) = S_\gamma \xi = T_\gamma \xi = z$, and $T_{\alpha^{-1}} T_\beta T_\gamma \xi \sim z$. By the remark above $T_\alpha T_{\alpha^{-1}} T_\beta T_\gamma \xi \sim T_\alpha z = T_\alpha T_\gamma \xi$ and $T_\beta T_\gamma \xi \sim T_\alpha T_\gamma \xi$ as was to be proved.

Given $x, y, z \in X$ with $x = T_\alpha \xi$, $y = T_\beta \xi$, and $z = T_\alpha T_\beta \xi$ we define $[x] \circ [y] = [z]$. If $x' \sim x$, $x' = T_{\alpha'} \xi$, $y' \sim y$, $y' = T_{\beta'} \xi$, and if $T_{\alpha'} T_{\beta'} \xi = z'$ exists, we claim $z' \sim z$. First $T_{\alpha'} T_{\beta'} \xi \sim T_{\alpha'} T_\beta \xi$ since $T_{\beta'} \xi \sim T_\beta \xi$. Then by Lemma 3.2.1, $T_{\alpha'} T_\beta \xi \sim T_\alpha T_\beta \xi$. By transitivity $z' \sim z$, and hence $[x] \circ [y] = [z]$ depends only on the classes $[x]$ and $[y]$ of x and y.

LEMMA 3.2.2. *If* $\alpha = \{\alpha_j\}_{j \in \Lambda}$ *and* $\beta = \{\beta_j\}_{j \in \Lambda}$ *are nets such that* $T_\alpha T_\beta \xi$ *and* $T_{\alpha\beta} \xi$ *exist where* $\alpha\beta = \{\alpha_j \beta_j\}_{j \in \Lambda}$, *then* $T_\alpha T_\beta \xi \sim T_{\alpha\beta} \xi$.

Proof. We claim that $T_{\beta^{-1}\alpha^{-1}} T_\alpha T_\beta \xi = \xi$ from which our assertion follows. Given a finite set N and $\epsilon > 0$ choose by Corollary 2.1.2 for $n = 3$ a finite set $M \supset N$ and $\delta > 0$ fulfilling the conclusion of that lemma. By our usual relative density argument choose subnets $\alpha_1 = \{\alpha_{j(i)}\}_{i \in \Lambda_1}$ and $\beta_1 = \{\beta_{j(i)}\}_{j \in \Lambda_1}$, such that $\beta_{j(i)} = r\tau_{j(i)}$ with $\tau_{j(i)} \in C_\delta(M)$. Assuming now that $T_{\beta^{-1}\alpha^{-1}} T_\alpha T_\beta \xi = \eta$ exists and $T_{\alpha_1^{-1}} T_{\alpha_1} \xi = \xi$ holds, choose i, then k, then l in Λ_1 so that

a) $r^{-1}\alpha_{j(i)}^{-1}\alpha_{j(k)} r \in C_\delta(M)$ and

b) $\max_{s,t \in N} |\xi(s\beta_{j(i)}^{-1}\alpha_{j(i)}^{-1}\alpha_{j(k)}\beta_{j(l)}, t) - \eta(s,t)| < \epsilon$.

By our choice of M and i and k we have that

$$\max_{s,t \in N} |\xi(s\beta_{j(i)}^{-1}\alpha_{j(i)}^{-1}\alpha_{j(k)}\beta_{j(l)}, t) - \xi(s,t)|$$
$$= \max_{s,t \in N} |\xi(s\tau_{j(i)}^{-1}r^{-1}\alpha_{j(i)}^{-1}\alpha_{j(k)} r\tau_{j(l)}, t) - \xi(s,t)| < \epsilon.$$

Letting ϵ tend to zero gives $\xi = \eta$. Thus if α_1, β_1 are nets as above for which $T_{\beta_1^{-1}\alpha_1^{-1}} T_{\alpha_1} T_{\beta_1} \xi = \eta$ and $T_{\alpha_1^{-1}} T_{\alpha_1} \xi = \xi$ hold, then $\xi = \eta$. Thus, as we have remarked before, this implies that $T_{\beta^{-1}\alpha^{-1}} T_\alpha T_\beta \xi = \xi$ and the lemma is proved.

Remark. The condition $T_\alpha T_\beta f = T_{\beta\alpha} f$ [1] has been called strict almost automorphy by Bochner. ([5].) Properly defined it is equivalent with almost periodicity (Bochner [5]). While $T_\alpha T_\beta = T_{\beta\alpha}$ does not hold for almost automorphic functions we do have the weaker statement that $T_\alpha T_\beta \sim T_{\beta\alpha}$.

THEOREM 3.2.1. *With the operation $[x] \circ [y]$ X_0 becomes a compact topological group. That is*

1) $([x] \circ [y]) \circ [z] = [x] \circ ([y] \circ [z])$ *for* $[x], [y], [z] \in X_0$.
2) $[x] \circ [\xi] = [\xi] \circ [x] = [x]$ *for all* $[x] \in X_0$.
3) *For each* $[x] \in X_0$ *there is a* $[y] = [x]^{-1}$ *such that*
$$[x] \circ [y] = [y] \circ [x] = [\xi].$$
4) *The mapping* $([x], [y]) \to [x] \circ [y]^{-1}$ *is continuous from* $X_0 \times X_0$ *to* X_0.

Proof. Let $x \in [x]$, $y \in [y]$, $z \in [z]$ be given by $x = T_\alpha \xi$, $y = T_\beta \xi$, $z = T_\gamma \xi$ with $\alpha = \{\alpha_i\}_{i \in \Lambda_1}$, $\beta = \{\beta_j\}_{j \in \Lambda_2}$, $\gamma = \{\gamma_k\}_{k \in \Lambda_3}$. If $T_\alpha T_\beta \xi$ exists there will clearly exist a net $\delta = \{\delta_l\}_{l \in \Lambda_4}$ where $\delta_l = \alpha_{i(l)} \beta_{j(l)}$ with $T_\delta \xi = T_\alpha T_\beta \xi$. Then $[T_\delta T_\gamma \xi] = ([x] \circ [y]) \circ [z]$. Assuming as we may that $\{\alpha_{i(l)}\}_{l \in \Lambda_4}$ and $\{\beta_{j(l)}\}_{l \in \Lambda_4}$ are subnets of α and β we have by the preceding lemma that $T_\delta T_\gamma \xi \sim T_\alpha T_\beta T_\gamma \xi$. But $[T_\alpha T_\beta T_\gamma \xi] = [x] \circ ([y] \circ [z])$ and 1) holds.

2) is trivial. For 3) let $[y] = [T_{\alpha^{-1}} \xi]$ if $T_\alpha \xi = x$. Then $[x] \circ [y] = [\xi] = [y] \circ [x]$ since ξ is almost automorphic.

Suppose now we have a net $([x], [y])_i$, $i \in \Lambda$ of elements of $X_0 \times X_0$ such that $\lim_{i \in \Lambda} ([x], [y])_i = ([x], [y])$ and $\lim_{i \in \Lambda} [x]_i \circ [y]_i^{-1} = [w]$. We choose $x_i \in [x]_i$ and $y_i \in [y]_i$ for each $i \in \Lambda$ and assume without relettering that Λ has been so refined that a) $\lim_{i \in \Lambda} x_i = x$, $\lim_{i \in \Lambda} y_i = y$ exist and b) if $x_i = T_{\alpha_i} \xi$, $y_i = T_{\beta_i} \xi$ then $w_i = T_{\alpha_i} T_{\beta_i^{-1}} \xi$ exists and $\lim_{i \in \Lambda} w_i = w$. Since π is continuous, $\pi x = [x]$, $\pi y = [y]$, and $\pi w = [w]$. We write $\alpha_i = \{\alpha_{i,j}\}_{j \in \Lambda_i}$, $\beta_i = \{\beta_{i,j}\}_{j \in \Gamma_i}$. For each finite set $N \subset G$ let an index $i \in \Lambda$ be chosen so that

c) $$\max_{s,t \in \Lambda} |x_i(s,t) - x(s,t)| < \frac{1}{|N|},$$

d) $$\max_{s,t \in \Lambda} |y_i(s,t) - y(s,t)| < \frac{1}{|N|},$$

[1] When dealing with functions $T_\alpha T_\beta = T_{\beta\alpha}$ not $T_{\alpha\beta}$ since $(\gamma\delta)f = \delta(\gamma f)$.

and

e) $$\max_{s,t\in N} | w_i(s,t) - w(s,t) | < \frac{1}{|N|}.$$

Writing $i = i(N)$ choose $\alpha_N = \alpha_{i,j}$, $\beta_N = \beta_{i,k}$ such that

c′) $$\max_{s,t\in N} | L_{\alpha_N}\xi \cdot (s,t) - x_i(s,t) | < \frac{1}{|N|},$$

d′) $$\max_{s,t\in N} | L_{\beta_N}\xi \cdot (s,t) - y_i(s,t) | < \frac{1}{|N|}, \text{ and}$$

e′) $$\max_{s,t\in N} | L_{\alpha_N\beta_N^{-1}}\xi \cdot (s,t) - w_i(s,t) | < \frac{1}{|N|}.$$

Letting $\alpha = \{\alpha_N\}$, $\beta = \{\beta_N\}$, and $\alpha\beta^{-1} = \{\alpha_N\beta_N^{-1}\}$ we have from c), d), e) and c′), d′), e′) that $T_\alpha\xi = x$, $T_\beta\xi = y$ and $T_{\alpha\beta^{-1}}\xi = w$. Assuming after a further refinement if necessary that $T_\alpha T_{\beta^{-1}}$ exists, we have from Lemma 3.2.2 that $T_\alpha T_{\beta^{-1}}\xi \sim T_{\alpha\beta^{-1}}\xi$ Thus $\pi w = [x] \circ [y]^{-1} = [w]$, and $[x] \circ [y]^{-1}$ is continuous from $X_0 \times X_0$ onto X_0. This completes the proof that X_0 is a topological group. (X_0 is compact as we earlier remarked.)

Paraphrasing Theorem 3.2.1 there is a compact group X_0 with two actions (G_R, X_0) and (G_L, X_0) of G on X_0 and a closed mapping π of X onto X_0 such that the diagrams

$$\begin{tikzcd} X \arrow[r, "G_R"] \arrow[d, "\pi"'] & X \arrow[d, "\pi"] \\ X_0 \arrow[r, "G_R"] & X_0 \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "G_L"] \arrow[d, "\pi"'] & X \arrow[d, "\pi"] \\ X_0 \arrow[r, "G_L"] & X_0 \end{tikzcd}$$

are commutative. By defintion of πx, $\pi^{-1}\pi x = \{x\}$ if and only if x is an almost automorphic point. (The set of almost automorphic points is of course dense in X.)

Remark. It can be shown that when G is countable, the set of points which are *not* almost automorphic is of the first category.

3.3. Applications. The mapping ϕ of G into X_0 given by $\phi(\gamma) = \pi L_\gamma\xi$ is a homomorphism of G onto a dense subgroup of X_0. The image G_0 of G in X_0 is isomorphic with G/H for some normal subgroup H of G. If $h \in C(X_0)$, the algebra of continuous complex valued functions on X_0, then h is almost periodic on X_0. ([13].) Thus $h_0(\gamma) = h(\pi L_\gamma\xi)$ is almost periodic on G (being almost periodic on G_0).

Theorem 3.1.1 implies that $x \in X$ is right almost automorphic if and

only if it is left almost automorphic. Suppose that f is such that $T_\alpha f$ is almost automorphic whenever the limit exists. Then for each $\gamma \in G$ the function $T_{\gamma\alpha} f$ will be almost automorphic if the limit exists. In X therefore $T_\alpha f$ will be right almost automorphic whenever the limit exists since for $s, t \in G$, $T_\alpha \xi \cdot (s,t) = T_{s\alpha} f \cdot (t)$. Then $T_\alpha \xi$ is left almost automorphic and $\pi^{-1}\pi T_\alpha \xi = T_\alpha \xi$. That is, the mapping $\pi\colon X \to X_0$ is one to one and hence a homeomorphism. Thus we have

THEOREM 3.3.1. *A function f on a group G is almost periodic if and only if it is almost automorphic and each function $T_\alpha f$ which arises from f is almost automorphic.*

Proof. If $T_\alpha f$ is almost automorphic whenever the limit exists then as has been observed above π is one to one and X is a compact group. Thus f is almost periodic on X and hence also on G. The other direction follows as in Theorem 2.3.1.

The remainder of this section will be devoted to characterizing the almost automorphic functions in terms of the almost periodic functions.

LEMMA 3.3.1. *Let f be a real valued almost automorphic function on G with X, X_0 as before. There exists a net $\{g_N\}$ of continuous functions on X_0 with index set the directed system of finite subsets of G such that*

i) $g_N(\pi L_x \xi) = f(x)$ *for* $x \in N' = N \cdot N = \{st\}_{s,t \in N}$.

ii) *With each $x \in N_0' = N_0 \cdot N_0$ where $N_0 \subseteq N$ is associated a closed neighborhood $U_x^{N_0}$ of $\pi L_x \xi$ such that*

$$\operatorname{Var} g_N\,|_{U_x^{N_0}} \leqq \operatorname{Var} f\,|_{\pi^{-1}U_x^{N_0}} \leqq \frac{1}{|N_0|}.$$

(By $\operatorname{Var} h\,|_V$ we mean $\sup\limits_{y,z \in V} |h(z) - h(y)|$.)

iii) $\max f = \max g_N$, $\min f = \min g_N$ *on their respective domains.*

Proof. We shall construct $\{g_N\}$ inductively on the cardinality $|N|$ of N. For $|N| = 0$, $N = \emptyset$ set $g_N = 0$. Suppose now that for each subset M of G with $|M| < n$ has been constructed a function g_M and set U_x^M, $x \in M$ satisfying i)-iii) above. Let N consist of n distinct elements. For each $N_0 \subset N$ there is a finite collection $U_x^{N_0}$ of closed neighborhoods of the point $\pi L_x \xi$, $x \in N_0'$. The totality of these as N_0 varies through the proper subsets of N is finite. For each $x \in N'$ we construct a closed neighborhood U_x^N of $\pi L_x \xi$ such that

a) if $\pi L_x\xi \notin U_x^{N_0}$, $N_0 \subset N$ then $U_z^{N_0} \cap U_x^N = \emptyset$,

b) $U_z^N \cap U_x^N = \emptyset$, $z, x \in N$, $z \neq x$, and

c) $\operatorname{Var} f|_{\pi^{-1}U_x^N} \leqq \frac{1}{|N|}$.

Denoting by $S(N)$ the set $\{U_x^{N_0}\}_{x \in N_0', N_0 \subset N}$ form the set $\mathfrak{F}_N$ consisting of all sets of the form $A = \bigcap_{i=1}^{k} U_{x_i}^{N_i}$ where $U_{x_i}^{N_i} \in S(N)$. $\mathfrak{F}_N$ is finite; $\mathfrak{F}_N = \{A_1, \cdots, A_l\}$. A set $A \in \mathfrak{F}_N$ will be called an atom if no set $B \in \mathfrak{F}_N$ is properly contained in A. For each $x \in N'$, $\pi L_x\xi$ is contained in one atom A_x, the intersection of those sets in $S(N)$ which contain $\pi L_x\xi$. For if $U_z^{N_0} \cap A \neq \emptyset$, then $U_z^{N_0} \cap U_x^N \neq \emptyset$, and by construction $\pi L_x\xi \in U_z^{N_0}$. Thus $U_z^{N_0} \cap A = A$, and A is an atom. Now g_N will be successively defined on the element of $\mathfrak{F}_N$ beginning with the atoms and finally extended to all of X_0. If A_x is the atom containing $\pi L_x\xi$, define $g_N \equiv f(x)$ on A_x. If A is an atom containing no points $\pi L_x\xi$, $x \in N'$, then define g_N on A to be constantly $f(y)$ for some $y \in \pi^{-1}A$, y being otherwise arbitrary. Note that if $A_1, \cdots, A_k$ are atoms, then $\operatorname{Var} g_N|_{A_1 \cup \cdots \cup A_k} \leqq \operatorname{Var} f|_{\pi^{-1}(A_1 \cup \cdots \cup A_k)}$ since on each A_j, g_N eakes on a value of f on $\pi^{-1}A_j$. Suppose now that g_N is defined and continuous on all sets $A \in \mathfrak{F}_N$ which properly contain at most $m-1$ set of $\mathfrak{F}_N$. Assume that d) $\operatorname{Var} g_N|_A \leqq \operatorname{Var} f|_{\pi^{-1}A}$, e) if $A_1, \cdots, A_k$ are k such sets, then $\operatorname{Var} g_N|_{A_1 \cup \cdots \cup A_k} \in \operatorname{Var} f|_{\pi^{-1}(A_1 \cup \cdots \cup A_k)}$ f) $\max f \geqq \max g_N \geqq \min g_N \geqq \min f$ on their respective domains. Let $A \in \mathfrak{F}_N$ contain properly m elements of $\mathfrak{F}_N$ numbering them $A_1, \cdots, A_m$. For each j g_N is already defined on A_j since A_j can contain properly at most $m-1$ elements of $\mathfrak{F}_N$. By the Tietze extension theorem g_N may be extended to A so that

$$\min_{\substack{1 \leqq i \leqq m \\ x \in A_i}} g_N(x) = \min_{x \in A} g_N(x) \quad \text{and} \quad \max_{\substack{1 \leqq i \leqq m \\ x \in A_i}} g_N(x) = \max_{x \in A} g_N(x).$$

Since $\operatorname{Var} g_N|_A = \max_{x \in A} g_N(x) - \min_{x \in A} g_N(x)$, it follows that

$$\operatorname{Var} g_N|_A \leqq \operatorname{Var} g_N|_{A_1 \cup \cdots \cup A_m} \leqq \operatorname{Var} f|_{\pi^{-1}(A_1 \cup \cdots \cup A_m)} \leqq \operatorname{Var} f|_{\pi^{-1}A}.$$

By the same token if $B_1, \cdots, B_j$ are sets of $\mathfrak{F}_N$ containing properly at most m elements, then $\operatorname{Var} g_N|_{B_1 \cup \cdots \cup B_j} \leqq \operatorname{Var} f|_{\pi^{-1}(B_1 \cup \cdots \cup B_j)}$. By induction g_N is defined on all sets $A \in \mathfrak{F}_N$ and satisfies d), e), and f) above. Since $U_x^N \in \mathfrak{F}_N$, g_N is defined on U_x^N and by d) $\operatorname{Var} g_N|_{U_x^N} \leqq \operatorname{Var} f|_{\pi^{-1}U_x^N} \leqq \frac{1}{|N|}$. Let g_N be defined on all of X_0 by making one more Tietze extension from $A_1 \cup \cdots \cup A_l$ to X_0. ($\mathfrak{F}_N = \{A_1, \cdots, A_l\}$.) Then g_N satisfies i)-iii) above and the lemma obtains by induction.

Definition 3.3.1. A net $\{f_N\}$ of almost automorphic functions on G shall be called *jointly almost automorphic* if it is uniformly bounded, and if for each finite set $M \subset G$ and $\epsilon > 0$ there exists a set $B_\epsilon = B_\epsilon(M)$ and an index N_0 such that if $N \geqq N_0$ then B_ϵ satisfies i)-iv) of Definition 2.1.2 for f_N, M, and ϵ.

THEOREM 3.3.2. *A function f on G is almost automorphic if and only if it is the pointwise limit of a jointly almost automorphic net of almost periodic functions.*

Proof. Suppose f is the pointwise limit of a jointly almost automorphic net $\{f_N\}$ where f_N is almost periodic on G. Then f is bounded since the net $\{f_N\}$ is uniformly bounded. Given a finite set $M \subset G$ and an $\epsilon > 0$ choose $B_{\epsilon/2}(M)$ and N_0 by Definition 3.3.1. Then since strict inequalities in the limit become at worst equalities it is clear that $B_{\epsilon/2}(M)$ satisfies the properties i)-iv) of Definition 2.1.2 of a set $B_\epsilon(M)$. (That is if

$$\max_{s,t \in M} | f_N(s\tau t) - f_N(st) | < \epsilon/2$$

for $N \geqq N_0$ then $\max_{s,t \in M} | f(s\tau t) - f(st) | \leqq \epsilon/2 < \epsilon$.)

Remark. We have used in this first part only the almost automorphy of almost periodic functions.

Suppose conversely that f is real and almost automorphic. Let $\{g_N\}$ be the net of continuous functions constructed in Lemma 3.3.1. Define $f_N(t) = g_N(\pi L_t\xi)$, $t \in G$. Then f_N is almost periodic on G for each N, and the net $\{f_N\}$ is uniformly bounded since $\inf f \leqq \inf f_N \leqq \sup f_N \leqq \sup f$ for each N. Also $\lim_N f_N(t) = f(t)$ for each $t \in G$ since $f_N(t) = f(t)$ as soon as $t \in N'$. Suppose there is given a finite set $M \subset G$ and $\epsilon > 0$. We are to produce a set $B_\epsilon(M)$ and index N_0 in accordance with Definition 3.3.1. Let first $M_1 \supset M$ be chosen so that $\dfrac{1}{|M_1|} < \epsilon$. Associated with M_1 are sets $\{U_{x_j}{}^{M_1}\}_{j=1,\cdots,l}$ which are closed neighborhoods of the points $\{\pi L_{x_j}\xi\}_{j=1,\cdots,l}$ in X_0. Writing $x_j = s_j t_j$ for each $j = 1, \cdots, l$ with $s_j, t_j \in M_1$, the (s,t)-th coordinate of $L_{x_j}\xi$ is given by $f(ss_jt_jt)$. Since $U_{x_j}{}^{M_1}$ is a neighborhood of $\pi L_{x_j}\xi$ there exists a set $M_0 \subset G$ and a $\delta > 0$ such that whenever

$$\max_{s,t \in M_0} | f(ss_j\tau t_j t) - f(ss_jt_jt) | < \delta,$$

then $\pi L_{s_j\tau t_j}\xi \in U_{x_j}{}^{M_1}, j = 1, \cdots, l$. Select by Corollary 2.1.2 for $n = 2$ a set B of such τ's with the property $\sigma, \tau \in B$ then $\pi L_{s_j\sigma^{\epsilon_1}\tau^{\epsilon_2}t_j}\xi \in U_{x_j}{}^{M_1}$, $j = 1, \cdots, l$, $\epsilon_i = 0$, 1, or -1. Define $B_\epsilon = B \cup B^{-1}$. If $N \geqq M_1$, and if $\tau \in B_\epsilon(M)$, then

$$\max_{s,t \in M} | f_N(s\tau t) - f_N(st) | \leqq \operatorname{Var} f\,|_{\pi^{-1}U_{x_j}{}^{M_1}} \leqq \frac{1}{|M_1|} < \epsilon.$$

Similarly if $\sigma, \tau \in B_\epsilon(M)$, $\max_{s,t \in M} |f_N(s\sigma\tau^{-1}t) - f_N(st)| < \epsilon$. Thus $B_\epsilon(M)$ satisfies the conditions of Definition 3.3.1 ($N_0 = M_1$), and $\{f_N\}$ is jointly almost automorphic. If f is complex valued, then $f = u + iv$ for certain real valued almost automorphic functions u and v. The above construction can be carried out simultaneously for u and v producing nets $\{u_N\}$ and $\{v_N\}$ such that $\{u_N + iv_N\}$ is almost automorphic. (For $x \in N'$ U_x^N is chosen so that $\max(\operatorname{Var} u|_{\pi^{-1}U_x^N}, \operatorname{Var} v|_{\pi^{-1}U_x^N}) < \frac{1}{|N|}$.)

3.4. General flows. While it is not our principal interest we remark briefly here on what can be said for a general flow. If G acts on a compact Hausdorff space X (say on the left) an almost automorphic point is defined by the condition $T_{\alpha^{-1}}T_\alpha\xi = \xi$ when the limit exists. It can be shown by essentially the same techniques as above that if there exists an almost automorphic point $\xi \in X$ whose orbit $\bigcup_{\gamma \in G} \gamma\xi$ is dense in X, then there exists a compact group Γ with closed subgroup Γ_0 such that G acts on the homogeneous space $\Gamma/\Gamma_0 = X_0$ and such that there is a closed mapping $\pi: X \to X_0$ which makes the diagram

$$\begin{array}{ccc} X & \xrightarrow{G} & X \\ \pi\downarrow & & \downarrow\pi \\ X_0 & \xrightarrow{G} & X_0 \end{array}$$

commutative. Furthermore $\pi^{-1}\pi x = \{x\}$ exactly when x is an almost automorphic point.

4. Continuous almost automorphy.

4.1. Continuity and countability assumptions. In the present section we shall consider (additive) groups G which are topological and

a) abelian c) σ-compact

b) locally compact d) first countable

Definition 4.1.1. A continuous function f on G will be called *continuous Bochner almost automorphic* if from every sequence $\{\alpha_n'\} = \alpha'$ of group elements may be extracted a subsequence $\alpha = \{\alpha_n\}$ such that $T_\alpha f$ exists and is continuous and $T_{\alpha^{-1}}T_\alpha f = f$.

Definition 4.1.2. A function f on G will be called *continuous Bohr almost automorphic* if it is Bohr almost automorphic and uniformly continuous.

LEMMA 4.1.1. *A continuous Bochner almost automorphic function is uniformly continuous.*

Proof. If f were not uniformly continuous on G, there would exist an $\epsilon > 0$ and sequences $\{t_n'\}$, $\{x_n'\}$ of group elements such that for each n $|f(x_n' + t_n') - f(x_n')| \geqq \epsilon$ and $\lim_{n\to\infty} t_n' = e$. For the latter we have used the first countability of G. Since f is continuous Bochner almost automorphic there exist subsequences $\{x_n\}$ and $\{t_n + x_n\}$ such that both $\lim_{n\to\infty} f(t + x_n) = g(x)$ and $\lim_{n\to\infty} f(t + t_n + x_n) = h(x)$ exist and are continuous with $|y(0) - h(0)| > \epsilon$. The set $N = \{y \mid |g(y) - h(y)| > 3\epsilon/4\}$ is open and contains e. For each n we define

$$A_n' = \{t \mid |f(x_m + t_m + t) - h(t)| \leqq \epsilon/4, |f(x_m + t) - g(t)| \leqq \epsilon/4, m \geqq n\}.$$

Since f, g, and h are by assumption continuous, the set A_n' is closed. Let $A_n = A_n' \cap N$. Then $\cup A_n = N$, and N being a Baire space (since N is locally compact and regular) we may apply the Baire category theorem to conclude that for n sufficiently large A_n contains an interior point x of N. We now choose m so large that 1) $x + t_m \in N$ and 2) $|g(x) - g(t_m + x)| < \epsilon/4$. Then by the triangle inequality

$$\begin{aligned}|g(x) - h(x)| \leqq |g(x) - g(t_m + x)| + |g(t_m + x) - f(x + t_m + x_m)| \\ + |f(x + t_m + x_m) - h(x)| < \epsilon/4 + \epsilon/4 + \epsilon/4 = 3/4.\end{aligned}$$

Thus $x \notin N$ contradicting our earlier assumption that $x \in N$ and proving our lemma.

THEOREM 4.1.1. *A function f is continuous Bochner almost automorphic if and only if it is continuous Bohr almost automorphic.*

Proof. If f is continuous Bohr almost automorphic, then f is Bochner almost automorphic by Theorem 2.2.1. We need only verify that sequences suffice and $T_\alpha f$ is continuous when it exists. Both of these last follow from the diagonal procedure applied to the two facts a) for K compact the functions γf, $\gamma \in G$ are equally uniformly continuous on K, and b) G is σ-compact. Conversely if f is continuous Bochner almost automorphic, it is uniformly continuous by the preceding lemma and Bohr almost automorphic by Theorem 2.2.1. This completes the proof.

For what follows we shall need a restatement of Corollary 2.1.2 in the context of continuous almost automorphy.

COROLLARY 2.1.2′. *If K is a compact set in G, and if $\epsilon > 0$ is given, then for any integer $n > 0$ there exists a compact set $K' \supset K$ and a $\delta > 0$*

such that if $\tau_1, \cdots, \tau_n \in C_\delta(K')$ *and if* ϵ_i *are numbers equal to* 0, 1, *or* -1, *then* $\sum_{i=1}^{n} \epsilon_i \tau_i \in C_\epsilon(K)$.

Proof. Let the compact set K be covered by finitely many translates $x_1 + u, \cdots, x_l + u$ where U is a neighborhood of e chosen for $\epsilon/3$ in the statement of uniform continuity. By Corollary 2.1.2 we choose a finite set M and $\delta > 0$ such that if $\tau_1, \cdots, \tau_n \in C_\delta(M)$, and $\epsilon_i = 0, 1$, or -1, then $\tau = \sum \epsilon_i \tau_i \in C_{\epsilon/3}(\{x_1, \cdots, x_l\})$. If $x \in K$, we have by the triangle inequality that

$$|f(x+\tau) - f(x)| \leqq |f(x+\tau) - f(x_j+\tau)| + |f(x_j+\tau) - f(x_j)| + |f(x_j) - f(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$$

choosing j so that $x \in U + x_j$. We let then $K' = M \cup K$. Note that our argument shows that $C_{\epsilon_0}(K_0)$ is relatively dense for a compact set K_0 and $\epsilon_0 > 0$.

4.2. The Banach algebra A_c. Theorem 1.2.1 remains true for the set A_c of continuous almost automorphic functions on G as one easily verifies. Thus A_c is a commutative C^* algebra with identity and by the commutative Gelfand-Naimark theorem, a $C(X)$. That is, A_c is isometrically isomorphic with the algebra of continuous complex valued functions on a compact Hausdorff space X, the maximal ideal space of A_c. ([11].) To each $t \in G$ corresponds a maximal ideal $M_t = \{f \in A_c \mid f(t) = 0\}$ and hence a point $\bar{t}$ of X. This natural mapping of G into X may not be one to one, but one easily proves the existence of a closed subgroup $K \subset G$ (namely the set of $t \in G$ such that $f(t) = f(e)$ for all $f \in A_c$) such that each $f \in A_c$ is constant on the cosets of K, and the induced mapping of G/K into X is one to one. Therefore we assume the mapping of G into X is one to one by first factoring out an appropriate subgroup. We shall not distinguish between points t of G and corresponding points of X, nor shall we distinguish between $f \in A_c$ and its representative in $C(X)$. In other words G will be considered a subset of X with each $f \in A_c$ extending from G to be continuous on X.

The operations $t + x$, $x \in X$, $t \in G$ considered as adjoint to the isomorphisms $f \to tf$ of A_c are homeomorphisms of X. Since G is commutative, it is well known that there exists a regular Borel probability measure μ on X invariant under G. That is, if C is a Borel subset of X then $\mu(t + C) = \mu(C)$, $t \in G$. The existence of μ follows for example from the Markov-Kakutani fixed point theorem. ([7]) If f is continuous on X, then for $t \in G$, $\int_X tf \cdot (x)\mu(dx) = \int_X f(x)\mu(dx)$. The measure μ will not in

general be unique however we fix it once and for all for the discussion. If χ_λ is a continuous character on G, then χ_λ is continuous almost automorphic and therefore defined on X. Define the Fourier coefficient

$$a_\lambda = \int_X \bar{\chi}_\lambda(x) f(x) \mu(dx).$$

4.3. Fourier analysis. Using an extension due to Følner of a result of Bogoliouboff we shall in this section prove that the Fourier series $f \sim \sum a_\lambda \chi_\lambda$ of a continuous almost automorphic function can be summed to the function by the Bochner-Fejer summation procedure. The summation will not of course be uniform but rather "jointly almost automorphic."

In preparation of the analysis let G be exhausted by an increasing sequence $\{K_j\}$ of compact sets, and associate with this sequence a sequence $\{\epsilon_j\}$ of positive numbers decreasing to zero. Apply Corollary 2.1.2′ for each j to K_j and ϵ_j with $n = 6$ to obtain sequences $\{K_j'\}$ of compact sets and $\{\delta_j\}$ of positive numbers such that if $\tau_1, \cdots, \tau_6 \in C_{\delta_j}(K_j')$, then $\tau = \sum_{i=1}^{6} \alpha_i \tau_i \in C_{\epsilon_j}(K_j)$ where $\alpha_i = 0, 1$, or -1. Assume also that $K_j' \subset K_{j+1}'$ and $\delta_{j+1} < \delta_j$ for each j. Define $B_j = C_{\delta_j}(K_j') \cup C_{\delta_j}(K_j')^{-1}$, and note that B_j is relatively dense and $B_j = B_j^{-1}$. Finally for each j let V_j be a neighborhood of e corresponding to ϵ_j in the statement of uniform continuity, and assume that $V_j \supset V_{j+1}$. According to Følner-Bogoliouboff ([10], [6]) if we let $E_j = B_j + B_j - B_j - B_j + V_j$, there exist continuous characters $\chi_{1,j}, \cdots, \chi_{q_j,j}$ on G such that if $t \in G$ satisfies $\operatorname{Re} \chi_{i,j}(t) \geqq 0$, $i = 1, \cdots, q_j$, then $t \in E_j$. Note that $E_j \supset E_{j+1}$ for each j. Thus by adding finitely many characters if necessary for each j we may assume that $\{\chi_{1,j}, \cdots, \chi_{q_j,j}\} \subset \{\chi_{1,j+1}, \cdots, \chi_{q_{j+1},j+1}\}$. For each j select integers $n_{1,j}, \cdots, n_{q_j,j}$ positive the only restriction being that $\sum_{i=1}^{q_j} \frac{2}{n_{i,j}} < \epsilon_j$. Defining

$$K_{i,j}(x) = \sum_{\nu=-n_{i,j}}^{n_{i,j}} \left(1 - \frac{|\nu|}{n_{i,j}}\right) \bar{\chi}_{i,j}{}^{\nu}(x) = \frac{1}{n_{i,j}} \frac{1 - \operatorname{Re} \bar{\chi}_{i,j}(x)^{n_{i,j}}}{1 - \operatorname{Re} \bar{\chi}_{i,j}(x)}$$

set up the function $K_j'(x) = \prod_{i=1}^{q_j} K_{i,j}(x)$. $K_j'(x)$ is almost periodic on G and therefore defined on X. The invariant mean for almost periodic functions being unique we must have that $\int_X K_j'(x) \mu(dx) = M_j \geqq 1$. The Bochner-Fejer kernel $K_j(x)$ is now given by $K_j(x) = K_j'(x)/M_j$ and satisfies i) $K_j(x) \geqq 0$ ii) $\int_X K_j(x) \mu(dx) = 1$ iii) $\int_{X-\bar{E}} K_j(x) \mu(dx) \leqq \sum_{i=1}^{q_j} \frac{2}{n_{i,j}} < \epsilon_j$. (For a proof of iii) cf. [9].)

Let

$$f_n(t) = \int_X f(x+t)K_n(x)\mu(dx) = \int_X f(x)K_n(x-t)\mu(dx). \tag{4.3.1}$$

The equality of the two integrals in (4.3.1) follows from the invariance of μ, and the second integral shows that $f_n(t)$ is a trigonometric polynomial. The functions $\{f_n(t)\}$ are uniformly bounded in norm by $\| f \|$. Furthermore if $s \in G$ is such that $\| f(t+s) - f(t) \| < \epsilon$, then $\| f_n(t+s) - f_n(t) \| \leqq \epsilon$ from i) and ii) above. Thus the functions $\{f_n(t)\}$ are equally uniformly continuous on G.

Definition 4.3.1. A sequence $\{f_n\}$ of continuous almost automorphic functions on G is *jointly continuous almost automorphic* if it is equally uniformly continuous and jointly almost automorphic.

THEOREM 4.3.1. *A function f on G is continuous almost automorphic if and only if it is the pointwise limit of a jointly continuous almost automorphic sequence of trigonometric polynomials. Furthermore the Fourier coefficients* $a_\lambda = \int_X \bar{\chi}_\lambda(x) f(x)\mu(dx)$ *where χ_λ runs through the continuous characters of G uniquely determine the function.*

Proof. If f is the pointwise limit of a jointly continuous almost automorphic sequence of trigonometric polynomials then f is almost automorphic by Theorem 3.3.2. The uniform continuity of f follows immediately from the joint uniform continuity of $\{f_n\}$. Hence f is continuous almost automorphic. Suppose conversely that f is continuous almost automorphic, and let $\{f_n\}$ be the sequence of trigonometric polynomials defined in (4.3.1). If $x \in \bar{E}_{n_0}$, then x can be approached by elements of the form $\tau_1 + \tau_2 - \tau_3 - \tau_4 + \tau$ where $\tau_j \in B_{n_0}$, $j = 1, \cdots, 4$ and $\tau \in V_{n_0}$. For these latter elements if $t \in K_{n_0}$, then

$$\begin{aligned} &| f(t + \tau_1 + \tau_2 - \tau_3 - \tau_4 + \tau) - f(t) | \\ &\quad \leqq | f(t + \tau_1 + \tau_2 - \tau_3 - \tau_4 + \tau) - f(t + \tau_1 + \tau_2 - \tau_3 - \tau_4) | \\ &\qquad + | f(t + \tau_1 + \tau_2 - \tau_3 - \tau_4) - f(t) | < \epsilon_{n_0} + \epsilon_{n_0} = 2\epsilon_{n_0}. \end{aligned}$$

Thus if $x \in \bar{E}_{n_0}$, $| f(t+x) - f(t) | \leqq 2\epsilon_{n_0}$. This gives us for $n \geqq n_0$ that

$$\begin{aligned} | f_n(t) - f(t) | &= \left| \int_X (f(t+x) - f(t)) K_n(x)\mu(dx) \right| \\ &\leqq \int_X | f(t+x) - f(t) | K_n(x)\mu(dx) \\ &\leqq \int_{\bar{X}-\bar{B}_{n_0}} | f(t+x) - f(t) | K_n(x)\mu(dx) \\ &\qquad\qquad + \int_{\bar{B}_{n_0}} | f(t+x) - f(t) | K_n(x)\mu(dx) \\ &\leqq 2 \| f \| \epsilon_n + 2\epsilon_{n_0} \end{aligned} \tag{4.3.2}$$

since $\int_{\bar{X}-\bar{E}_{n_0}} K_n(x)\mu(dx) \leqq \int_{\bar{X}-\bar{E}_n} K_n(x)\mu(dx) < \epsilon_n$. Equation (4.3.2) implies that $f_n(t)$ converges pointwise to f. A similar argument shows that if $n \geqq n_0$ and $\sigma, \tau \in B_{n_0}$, then $|f_n(t+\alpha_1\sigma+\alpha_2\tau) - f_n(t)| < 2\|f\|\epsilon_n + 2\epsilon_{n_0}$ for $t \in K_{n_0}$ and $\alpha_1, \alpha_2 = 0, 1$, or -1. Thus $\{f_n\}$ is jointly almost automorphic, and as remarked earlier $\{f_n\}$ is equally uniformly continuous. If $\int_X \bar{\chi}_\lambda(x) f(x)\mu(dx) = 0$ for each continuous character χ_λ, then $f_n(t) = 0$ for all n, and hence $f = 0$. In other words f is uniquely determined by its Fourier coefficients, and the theorem is proved.

4.4. Some special series. In what follows we assume $G = \boldsymbol{R}$ with its usual topology. The continuous characters of $\boldsymbol{R}$ are functions $\chi_\lambda = \exp(i\lambda t)$ for each real λ. The following two theorems can be proved by essentially the same arguments as in [1], pp. 51-52.

THEOREM 4.4.1. *If f is continuous almost automorphic on $\boldsymbol{R}$ and if the set of λ's such that $a_\lambda = \int_X f(x)\bar{\chi}_\lambda(x)\mu(dx) \neq 0$ is linearly independent over the rational numbers, then f is almost periodic and $\sum_\lambda |a_\lambda| < \infty$.*

THEOREM 4.4.2. *If f is continuous almost automorphic, and if $a_\lambda \geqq 0$ for each λ, then f is almost periodic, and $\sum a_\lambda < \infty$.*

Less obvious perhaps is the following analogue of a theorem of Favard for almost periodic functions.

THEOREM 4.4.3. *Let f be continuous almost automorphic on $\boldsymbol{R}$, and suppose that $f(t)$ has no nonzero Fourier coefficients in the interval $(-\delta, \delta)$, $\delta > 0$. Then $F(s) = \int_0^s f(t)\,dt$ is bounded and almost automorphic.*

Proof. Initially let us assume that f also has a bounded uniformly continuous derivative. The argument of [1], p. 6 remains unchanged to show that f' is continuous almost automorphic. By [3] § 23 there exists a function $K(t) \in L^1(-\infty, \infty)$ such that $\int_{-\infty}^{\infty} \exp(-i\lambda t)K(t)\,dt = \hat{K}(t) = \frac{1}{i\lambda}$ for $|\lambda| \geqq \delta$. Define $h(x) = \int_{-\infty}^{\infty} f(x-t)K(t)\,dt$. Both $h(x)$ and $h'(x)$ are almost automorphic as is easily verified. (For example the mapping $t \to f(x-t)$ is continuous from $\boldsymbol{R}$ to A_c and $h(x)$ is just the integral of this mapping with respect to a finite measure.)

By Fubini's theorem

$$\int_X \bar{\chi}_\lambda(x)h(x)\mu(dx) = \int_X \bar{\chi}_\lambda(x)\{\int_{-\infty}^{\infty} f(x-t)K(t)\,dt\}\mu(dx)$$

$$= \int_{-\infty}^{\infty} K(t) \int_X \bar{\chi}_\lambda(x) f(x-t)\mu,dx)\,dt$$

$$= \int_{-\infty}^{\infty} \bar{\chi}_\lambda(t)K(t)a_\lambda dt = \frac{a_\lambda}{i\lambda} \text{ if } |\lambda| \geqq \delta, = 0 \text{ if } |\lambda| < \delta.$$

We claim that $h(x) = h(0) + \int_0^X f(t)dt = h(0) + F(x)$. Since h is bounded and almost automorphic, it will then follow that F has the same property. Since the derivative of f is uniformly continuous the expressions $\dfrac{f(x+\epsilon)-f(x)}{\epsilon}$ converge uniformly in x to $f'(x)$, and so

$$\int_X \bar{\chi}_\lambda(x)\Big[\frac{f(x+\epsilon)-f(x)}{\epsilon}\Big]\mu(dx) = \frac{\chi_\lambda(\epsilon)-1}{\epsilon}a_\lambda$$

converges to $\int_X \bar{\chi}_\lambda(x)f'(x)\mu(dx)$. That is the Fourier coefficients of f' are $i\lambda\, a_\lambda$. By the same argument as above we conclude that the Fourier coefficient of $h'(x) = \int_{-\infty}^{\infty} f'(x-t)K(t)\,dt$ are a_λ. Thus h' and f have the same Fourier coefficients and must therefore be equal. Hence $h(x) = h(0) + F(x)$ as claimed. The general case now follows by a "smoothing" argument. Let $\{\phi_n(t)\}$ be a sequence of functions in $L^1(-\infty,\infty)$ with two bounded derivatives which are also L' such that a) $\int_{-\infty}^{\infty} \phi_n(t)\,dt = 1$, b) $\phi_n(t) \geqq 0$, and c) for each $\epsilon > 0$ $\lim_{n\to\infty} \int_{-\epsilon}^{\epsilon} \phi_n(t)\,dt = 1$. Then by a standard argument $f_n(x) = \int_{-\infty}^{\infty} \phi_n(x-t)f(t)\,dt$ is almost automorphic for each n and converges uniformly to $f(x)$. The functions f_n each have a uniformly continuous first derivative and no nonzero Fourier coefficients in $(-\delta,\delta)$. By the first part of our proof $h_n(x) = \int_{-\infty}^{\infty} f_n(x-t)K(t)\,dt$ is also given by

$$h_n(x) = h_n(0) + F_n(x) = h_n(0) + \int_0^x f_n(t)\,dt.$$

Letting $n\to\infty$ we have $h(x) = h(0) + \int_0^x f(t)\,dt$, and our theorem obtains.

By the same techniques one can prove

THEOREM 4.4.4. *Let* $\sum_{j=0}^{n} a_j \frac{d^j}{dx^j}$ *be an ordinary differential operator with constant coefficients, and let* $\lambda_1, \cdots, \lambda_k$ *be the real roots of the equation* $\sum_{j=0}^{n} a_j (ix)^j = 0$. *If* f *is continuous almost automorphic and if* f *has no non-zero Fourier coefficients in a neighborhood of each* λ_j, $j = 1, \cdots, K$, *then there exists a bounded almost automorphic solution to the equation* $\sum_{j=0}^{n} a_j \frac{d^j F}{dx^j} = f$.

5. Applications to differential equations.

Theorem 3.3.1 appears to serve as a natural bridge between the recent approach ([5]) of Bochner and the classical approach of Favard ([8]) to certain theorems which guarantee the existence of almost periodic solutions to systems of differential equations.

We will consider with Favard systems

$$(S_t) \qquad \frac{dx_i}{dt} = \sum_{j=1}^{n} f_{i,j}(t) x_j(t) + g_i(t), \qquad i = 1, \cdots, n$$

where for each i, j $f_{i,j}(t)$ and $g_i(t)$ are Bohr almost periodic on $\boldsymbol{R}$. If $f_{i,j}^*(t) = \lim_{n\to\infty} f_{i,j}(t + t_n)$ and $g_i^*(t) = \lim_{n\to\infty} g_i(t + t_n)$, then we consider also the system

$$(S_t^*) \qquad \frac{dx_i}{dt} = \sum_{j=1}^{n} f_{i,j}^*(t) x_j(t) + g_i^*(t), \qquad i = 1, \cdots, n.$$

The totality of systems (S_t^*) is written $H(S_t)$, and if we denote the homogeneous systems associated with (S_t), (S_t^*) by (Σ_t), (Σ_t^*) respectively, then $H(\Sigma_t)$ will be the totality of systems (Σ_t^*). If $(x_1(t), \cdots, x_n(t)) = x(t)$ is a vector valued function on $\boldsymbol{R}$, define the absolute value $|x(t)|$ to be $(\sum_{i=1}^{n} x_i(t)^2)^{\frac{1}{2}}$ for each t.

The second theorem of Favard states

THEOREM (*Favard*). *If none of the systems contained in* $H(\Sigma_t)$ *admits a nontrivial solution* $x(t)$ *such that* $\inf |x(t)| = 0$, *and if* (S_t) *has a bounded solution, then* (S_t) *has an almost periodic solution.*

Proof. By the first part of Favard's argument [8] there is a solution $x_0(t)$ to (S_t) such that $\| x_0 \| = \sup_{-\infty<t<\infty} |x(t)|$ is a minimum. Favard shows that the assumption on (Σ_t) implies that x_0 is unique. Both $x_0(t)$ and $x_0'(t)$ are uniformly continuous so that given a sequence $\{\alpha_n'\} \subset \boldsymbol{R}$ there exists a subsequence $\alpha = \{\alpha_n\}$ such that $T_{-\alpha}T_\alpha x_0 = y_0$ and $T_{-\alpha}T_\alpha x_0' = y_0'$ exist.

Furthermore we may assume since $f_{i,j}$ and g_i are almost automorphic that $T_{-\alpha}T_\alpha f_{i,j} = f_{i,j}$ and $T_{-\alpha}T_\alpha g_i = g_i$. Clearly y_0 is a solution to (S_t) and $\| y_0 \| \leqq \| x_0 \|$. Since $\| x_0 \|$ is a minimum and x_0 is the unique element with norm $\| x_0 \|$, it follows that $x_0 = y_0$, and therefore x_0 is continuous almost automorphic. Suppose now that α is a sequence such that $T_\alpha x_0$ exists, and assume as we may that $T_\alpha f_{i,j} = f_{i,j}^*$ and $T_\alpha g_i = g_i^*$ exist for all i, j. Then $T_\alpha x_0$ is a solution to one of the systems (S_t^*) of $H(S_t)$. We claim that $\| T_\alpha x_0 \|$ is a minimum. For if w_0 is a solution to (S_t^*) with $\| w_0 \| < \| T_\alpha x_0 \|$, then for some subsequence α' of α, $T_{-\alpha'} w_0$ exists and satisfies S_t. But clearly $\| T_{-\alpha'} w_0 \| < \| T_{-\alpha} T_\alpha x_0 \| = \| x_0 \|$. Therefore $T_\alpha x_0$ has minimum norm, and by the assumption on $\Sigma_t^* \in H(\Sigma_t)$ $T_\alpha x_0$ is unique. Hence by the first part of the argument $T_\alpha x_0$ is continuous almost automorphic, and by Theorem 3.3.1 x_0 is almost periodic.

Remark. Bochner has recently in [5] proved a theorem related to the theorems of Favard, and his proof after which ours is modeled likewise proceeds in two steps. He first verifies the almost automorphy of a certain solution X to (S_t), and then using the hypothesis on the translated systems he proves X is "strictly almost automorphic," a property equivalent with almost periodicity.

6. Examples.

6.1. Examples on $\boldsymbol{Z}$. Given an infinite sequence of subgroups of $\boldsymbol{Z}$ $G_0, G_1, \cdots$ with $G_0 = \boldsymbol{Z}$ and G_{k+1} contained properly in G_k it was shown in [18] that there is a sequence $\{a_k\}$ of integers such that if $A_k = a_k + G_k$ for each k, then i) $A_k \cap A_l = \emptyset$, and ii) $\bigcup_{k=1}^{\infty} A_k = \boldsymbol{Z}$. For each $k = 1, 2, \cdots$ let f_k be an almost automorphic function on G_k, and assume $\| f_k \| \leqq M < \infty$ for all k. A function f is defined on $\boldsymbol{Z}$ by setting $f(s) = f_k(s')$ where $s = a_k + s'$, $s' \in G_k$. Properties i) and ii) above imply that f is well defined on $\boldsymbol{Z}$.

THEOREM 6.1.1. *The function f defined above is almost automorphic.*

Proof. f is bounded since $\sup_k \| f_k \| = \| f \| < \infty$. If $s \in \boldsymbol{Z}$, then $s = a_k + s'$ with $s' \in G_k$ for some k. Given $\epsilon > 0$ there is since f_k is almost automorphic a set $B_\epsilon(s') \subset G_k$ satisfying i)-iv) of Definition 2.1.2 for f_k, G_k. $B_\epsilon(s')$ may equally well be thought of as a subset of $\boldsymbol{Z}$, and clearly $B_\epsilon(s) = B_\epsilon(s')$ continues to satisfy the requirements of Definition 2.1.2.

Remark. In [18] it was observed that f need not be almost periodic

even when f_k is periodic on G_k for each k. We do not have condition on $\{f_k\}$ that f be almost periodic, but it can be shown by arguing along the lines of [16] that if f_k is a constant function equal to c_k for each k, then $\lim_{k\to\infty} c_k = c$ existing is necessary and sufficient for f to be almost periodic.

6.2. Analytic almost automorphic functions.

LEMMA 6.2.1. *Let $f(n)$ be a bounded function on $\mathbf{Z}$, and define $F_m(z) = (\frac{\sin \pi z}{\pi})^m \sum_{n=-\infty}^{\infty} \frac{f(n)}{(z-n)^m}$ for an integer $m \geqq 2$. For $x \in \mathbf{Z}$ $F_m(x) = f(x)$, and $F_m(x)$ is uniformly bounded in horizontal strips $-a \leqq \operatorname{Im} z \leqq a$, $a > 0$.*

The proof of Lemma 6.2.1 is not difficult and will be omitted.

THEOREM 6.2.1. *$F_m(z)$ is almost automorphic in horizontal strips if and only if $f(n)$ is almost automorphic on $\mathbf{Z}$. $F_m(z)$ is almost periodic if and only if $f(n)$ is almost periodic on $\mathbf{Z}$.*

Proof. Since $F_m(z) = f(z)$ for integers z, it is clear that f must be almost automorphic or almost periodic if F_m is to be almost automorphic or almost periodic. Suppose conversely $f(n)$ is almost automorphic or almost periodic. Since $F_m(z)$ is uniformly bounded in horizontal strips, $F_m(z)$ is uniformly continuous in horizontal strips. (Cf. [1].) Therefore from a given sequence $\{x_n'\}$ of real numbers we may extract a subsequence $\{x_n\}$ such that i)

$$\lim_{j\to\infty} \lim_{i\to\infty} F_m(z + x_i - x_j) = H_m(z)$$

exists uniformly on compact subsets of the plane, ii) $x_i = l_i + r_i$ with $l_i \in \mathbf{Z}$, $0 \leqq r_i < 1$, and $\lim_{i\to\infty} r_i = r$ exists, and iii) $\lim_{i\to\infty} \lim_{j\to\infty} f(n + l_i - l_j) = f(n)$ holds for $n \in \mathbf{Z}$. For each i, j we write

$$F_m(z + x_i - x_j) = \frac{\sin \pi^m (z + x_i - x_j)}{\pi^m} \sum_{n=-\infty}^{\infty} \frac{f(n + l_j - l_i)}{(z - n + r_i - r_j)^m}.$$

Taking the limit on i then j and applying i)-iii) above gives $H_m = F_m$, and F_m is almost automorphic in strips. If $f(n)$ is almost periodic, and if $\{x_n\}$ is such that a) $\lim_{n\to\infty} F_m(z + x_n) = G_m(z)$ exists, b) $x_n = l_n + r_n$, $l_n \in \mathbf{Z}$, $0 \leqq r_n < 1$, c) $\lim_{n\to\infty} f(s - l_n) = g(s)$, $s \in \mathbf{Z}$ holds, and d) $\lim_{n\to\infty} r_n = r$, then $\lim_{n\to\infty} F_m(z + x_n) = (\frac{\sin \pi (z + r)}{\pi})^m \sum_{s=-\infty}^{\infty} \frac{g(s)}{(z - s + r)^m} = G_m(z)$, and the same

argument as before shows that $G_m(z)$ is almost automorphic. By Theorem 3.3.7 $F_m(z)$ is almost periodic completing the proof.

REFERENCES.

[1] A. S. Besicovitch, *Almost Periodic Functions*, Dover, New York, 1954.

[2] S. Bochner, "Curvature and Betti numbers in real and complex vector bundles," *Universitá e Politecnico di Toorino, Rendiconti del seminario matematico*, vol. 15 (1955-1956), pp. 225-254.

[3] ———, *Lectures on Fourier Integrals*, Princeton University Press, Princeton, 1959.

[4] ———, "Uniform convergence of monotone sequences of functions," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 47 (1961), pp. 582-585.

[5] ———, "A new approach to almost periodicity," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 48 (1962), pp. 2039-2043.

[6] N. Bogoliouboff, "Sur quelques propriétés arithmétiques des presques periodes," *Ann. Chaire Phys. Math.*, Kiev: 4 (1939), pp. 195-205.

[7] N. Dunford and J. T. Schwartz, *Linear Operators*, Interscience, New York, 1958.

[8] J. Favard, "Sur les équations différentielles linéaires á coefficients presque-périodiques," *Acta Mathematica*, vol. 51 (1927), pp. 31-81.

[9] E. Følner, "A proof of the main theorem for almost periodic functions in an abelian group," *Annals of Mathematics* (2), vol. 50 (1949), pp. 559-569.

[10] ———, "Generalization of a theorem of Bogoliouboff to topological abelian groups," *Mathematica Scandinavica*, vol. 2 (1954), pp. 5-19.

[11] J. L. Kelley, *General Topology*, Van Nostrand, Princeton, 1955.

[12] B. M. Levitan, "Some questions of the theory of almost periodic functions," Uspehi Matematiceskih Nauk SSSR (N. S.), vol. 2, no. 6 (22) (1947), pp. 174-214.

[13] L. H. Loomis, *Abstract Harmonic Analysis*, Van Nostrand, Princeton, 1953.

[14] W. Maak, *Fastperiodische Funktionen*, Springer, Berlin, 1950.

[15] J. von Neumann, "Almost periodic functions in a group, I," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 445-492.

[16] O. Toeplitz, "Ein beispeil zur theorie der fastperiodischen funktionen," *Mathematische Annalen*, vol. 98 (1927), pp. 281-295.

[17] W. A. Veech, *Almost Automorphic Functions, Thesis*, Princeton, March 1963.

[18] ———, "Almost automorphic functions," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 49 (1963), pp. 462-464.

FAMILIES OF LINEAR OPERATORS DEPENDING UPON A PARAMETER.

By Felix E. Browder.

Introduction. Let H_1 and H_2 be two Hilbert spaces, M a differentiable or analytic manifold, $\{T_t; t \in M\}$ a family of closed linear operators from H_1 to H_2 parametrized by t in M. Suppose that $R(T_t)$, the range of T_t, is closed in H_2 and that the domain of T_t is a linear subset W of H_1 which is independent of t in M, where W carries a stronger Hilbert space structure. Let $N(T_t)$ be the null space of T_t, and let P_t be the orthogonal projection of H_1 on $N(T_t)$, Q_t the orthogonal projection of H_2 on $R(T_t)$. The generalized inverse G_t of T_t is the bounded linear operator from H_2 to H_1 defined by the condition $u = G_t v$ if and only if $T_t u = Q_t v$, $P_t u = 0$ (i.e. intuitively $G_t = (I - P_t) T_t^{-1} Q_t$).

There have been a number of studies in recent years of the problem of deriving the functional properties in t of P_t, Q_t, and G_t in terms of the corresponding properties of T_t. These studies in connection with parametrized families of elliptic differential operators and elliptic boundary value problems on vector bundles have been carried out by Nirenberg in Kodaira-Spencer [4], by the writer in Browder [1], and by Hörmander in Weil [5], Appendix III and Hörmander [3], Theorem 10. 5. 3. The problem can be formulated more precisely in the following way: Let F be a class of functions from manifolds M to $L(H, H')$, the bounded linear operators between pairs of Hilbert spaces H and H'. We suppose that $t \to T_t$ is a function of class F from M to $L(W, H_2)$ and ask whether the functions $t \to P_t: M \to L(H_1, W)$ $t \to Q_t: M \to L(H_2, H_2)$, and $t \to G_t: M \to L(H_2, W)$ are all of class F.

It is our object in the present paper to apply the methods of [1] to obtain a very general result in this direction which includes all preceding results as special cases and with great generality both in the class of linear operators T_t allowed and the class F of functional dependence. In [3], [4], and [5], Fredholm operators T_t are considered with $\dim N(T_t)$ constant in t with F either C^r or C^∞. In [1], the class of operators T_t was essentially those with $N(T_t) = \{0\}$ while F was the class of real analytic functions. In the present paper, we consider operators T_t under the sole restriction that

Received July 17, 1964.

$\dim(N(T_t)) < +\infty$ and independent of t in M, and impose the following conditions on the class F of functional dependence:

Definition. F is said to be a permissible class if it satisfies all of the following conditions:

(1) F is determined by local conditions on M on the operator valued function T on M. F includes all constant functions.

(2) Every function T in F is continuous from M to the norm topology on $L(H,H')$.

(3) If A, B, $C \in F$ with $A: M \to L(H,H')$, $B: M \to L(H,H')$, and $C: M \to L(H',H'')$, then the functions $t \to A_t + B_t: M \to (H,H')$, $t \to C_t B_t$: $M \to L(H,H')$, and $t \to A_t^*: M \to L(H',H)$ must all lie in F.

(4) If $A: M \to L(H,H')$ belongs to F and if A_t for each t in M is an isomorphism of H onto H', then $t \to (A_t)^{-1}: M \to L(H',H)$ also lies in F.

Simple examples of permissible classes F are C^r for $r \geqq 0$, C^∞, real analytic functions, and various Gevrey classes.

The basic result of the present paper is the following:

THEOREM. *Let M be a manifold, F a permissible class of operator-valued functions, $\{T_t: t \in M\}$ a family of closed linear operators with closed range from one Hilbert space H_1 to a second Hilbert space H_2. Suppose that for all t in M, the domain of T_t is W, where W is a linear subset of H_1 independent of t which carries a Hilbert space structure such that the injection of W into H_1 is continuous. Suppose further that the function $t \to T_t: M \to L(W, H_2)$ lies in the class F and that $\dim N(T_t)$ is finite and independent of t on M. Then if P_t is the orthogonal projection of H_1 on $N(T_t)$, Q_t the orthogonal projection of H_2 on $R(T_t)$, G_t the generalized inverse of T_t, we have that the functions*

$$t \to P_t: M \to L(H_1, W),$$
$$t \to Q_t: M \to L(H_2, H_2),$$
$$t \to G_t: M \to L(H_2, W),$$

all lie in the class F.

The proof of the Theorem is contained in the sequence of Lemmas which follows:

LEMMA 1. *Let R_0 be a closed subspace of the Hilbert space H, $t \to S_t$:*

$M \to L(R_0, H)$ a function in the class F, such that for a given t_0 in M, $S_{t_0} = J$ where J is the injection map of R_0 into H. Then:

(a) *For t in some neighborhood D of t_0 in M, $R(S_t)$ is closed in H. If R_0 has finite codimension in H,* $\operatorname{codim} R(S_t) = \operatorname{codim} R_0 < +\infty$ *for t near t_0.*

(b) *If E_t is the orthogonal projection of H on $R(S_t)$ for t in D, then $t \to E_t\colon D_1 \to L(H,H)$ lies in F for some neighborhood D_1 of t_0 in M.*

(c) *If $t \to L_t\colon D \to L(H,H)$ is a function in F such that $R(L_t)$ is contained in $R(S_t)$ for t in D, then for t near t_0, the equation $S_t U_t = L_t$ has an unique solution U_t in $L(H, R_0)$ and the function $t \to U_t\colon D_2 \to L(H, R_0)$ lies in F for some neighborhood D_2 of t_0 in M.*

Proof of Lemma 1. *Proof of* (a). Since S is continuous from M to the norm topology on $L(R_0, H)$, there exists a neighborhood D of t_0 in M such that for t in D, $u \in R_0$, we have $\| S_t u - u \| \leqq \frac{1}{2} \| u \|$. Hence $\| S_t u \| \geqq \frac{1}{2} \| u \|$, so that S_t is one-to-one and has a bounded inverse on its range for t in D. For such t, $R(S_t)$ is therefore closed in H_2.

If R_0 has finite codimension r in H, then $R_1 = H - R_0$ is of dimension r. We extend S_t to an element of $L(H,H)$ by setting $S_t u = u$ for u in R_1. For this extended mapping and t in D, we have $\| S_t - I \| \leqq \frac{1}{2}$. Hence S_t maps H one-to-one onto H, R_1 is a closed complement for $R(S_t)$ in H, and $\operatorname{codim} R(S_t) = r = \operatorname{codim} R_0 < +\infty$.

Proof of (b). It suffices to show that $t \to E_t\colon D_1 \to L(H,H)$ lies in the class F for some neighborhood D_1 of t_0. For t in D as above, $v \in H$, there exists an unique element u_t of R_0 such that $S_t u_t = E_t v$. Since the function $S\colon M \to L(R_0, H)$ is of class F, it suffices to show that the function $t \to U_t\colon D \to L(H, R_0)$ is of class F where $U_t v = u_t$. Since $S_t u_t = E_t v$, we have

$$(v - S_t u_t, S_t w) = 0$$

for all w in R_0. Let $S_t^* \in L(H, R_0)$ be the adjoint map to S_t. Since $S_t = J + S'_t$, where $\| S'_t \| \to 0$ as $t \to t_0$, we see that

$$S_t^* = J^* + (S'_t)^* = E_0 + S''_t$$

where $E_0 = J^*$ is the orthogonal projection map of H on R_0 while $\| S''_t \| = \| S'_t \| \to 0$ as $t \to t_0$. We see from the above that

$$S_t^* S_t u_t = S_t^* v$$

and

$$S_t^* S_t u_t = (E_0 + S''_t)(J + S'_t) u_t = u_t + K_t u_t,$$

where, if we choose D_1 sufficiently small, $\| K_t \| < 1$. Hence for all t in D_1, $(I + K_t) : R_0 \to R_0$ is an isomorphism. Since the function $t \to (I + K_t) : D_1 \to L(R_0, R_0)$ lies in F, if we denote $(I + K_t)^{-1} : R_0 \to R_0$ by X_t, the function $t \to X_t : D_1 \to L(R_0, R_0)$ also lies in F. Since $E_t v = S_t X_t S_t^* v$, $E_t = S_t X_t S_t^*$, and the function $t \to E_t : D_1 \to L(H, H)$ lies in the class F.

Proof of (c). $S_t U_t = L_t$ if and only if $S_t^* S_t U_t = S_t^* L_t$ if and only if $U_t = X_t S_t^* L_t$.

LEMMA 2. *Let W and H_2 be two Hilbert spaces, $\{T_t; t \in M\}$ a family of bounded linear operators from W to H_2 such that the function $t \to T_t : M \to L(W, H_2)$ lies in the class F. Suppose that* $\dim(N(T_t))$ *is finite and constant in t on M. Let P'_t be the orthogonal projection map of W on $N(T_t)$ and suppose that $R(T_t)$ is closed in H_2 for each t in M.*

Then the function $t \to P'_t : M \to L(W, W)$ is of class F.

Proof of Lemma 2. Let $T_t^* : H_2 \to W$ be the adjoint of T_t. Then $t \to T_t^* : M \to L(H_2, W)$ is a function of class F, $R(T_t^*)$ is closed in W for each t in M, and $\operatorname{codim} R(T_t^*)$ is finite and constant in t on M. Let Q'_t be the orthogonal projection map of W on $R(T_t^*)$. Then $P'_t = I - Q'_t$, and it suffices to prove that for each given t_0 in M, there exists a neighborhood D of t_0 such that $t \to Q'_t : D \to L(W, W)$ is a function of class F.

Let $R_0 = R(T_{t_0}^*)$, $R_t = R(T_t^*)$, $N_0 = N(T_{t_0}^*)$, $C_0 = H_2 \ominus N_0$. The map $T_{t_0}^*$ is an isomorphism of C_0 on R_0, whose inverse we denote by Z_0. We form the mapping $S_t : R_0 \to W$ by setting

$$S_t = T_t^* Z_0.$$

Then $t \to S_t : M \to L(R_0, W)$ is a function of class F and S_{t_0} is the injection map J of R_0 into W. We apply Lemma 1 to this mapping and assert that $R(S_t)$ is closed in W for t near t_0, $\operatorname{codim} R(S_t) = \operatorname{codim} R_0$, and if E_t is the orthogonal projection of W on $R(S_t)$, then the function $t \to E_t : D \to L(W, W)$ is a function of class F for some neighborhood D of t_0. However, $R(S_t) \subset R_t$ and $\operatorname{codim} R(S_t) = \operatorname{codim} R_0 = \operatorname{codim} R_t$. Hence $R(S_t) = R_t$ for t in D, $Q'_t = E_t$, and the Lemma is proved.

LEMMA 3. *Let H_1 and H_2 be two Hilbert spaces, $\{T_t : t \in M\}$ a family of closed linear operators from H_1 to H_2 with closed ranges in H_2 and with*

domain W, where W is a Hilbert space independent of t which is contained as a linear subset of H_1 with a stronger Hilbert space structure. Suppose that $\dim N(T_t)$ *is finite and independent of t in M. Let P_t be the orthogonal projection of H_1 on $N(T_t)$, and suppose that the function*

$$t \to T_t \colon M \to L(W, H_2)$$

is of class F.

Then the function $t \to P_t \colon M \to L(H_2, W)$ is of class F.

Proof of Lemma 3. If we consider T_t as an element of $L(W, H_2)$, the family $\{T_t \colon t \in M\}$ satisfies the hypotheses of Lemma 2. Hence if Q'_t is the orthogonal projection of W on $N(T_t)$, the function $t \to Q'_t \colon M \to L(W, W)$ is of class F. Let S_t be the restriction of Q'_t to $N_0 = N(T_{t_0})$, J_0 the injection map of N_0 into W. Then $t \to S_t \colon M \to L(N_0, W)$ is of class F and hence so is the function $t \to S_t \colon M \to L(N_0, H_1)$. Since $S_{t_0} = J_0$, the hypotheses of Lemma 1 are satisfied for the family $\{S_t\}$, and hence if E_t is the orthogonal projection of H_1 on $R(S_t)$, the function $t \to E_t \colon D \to L(H_1, H_1)$ is of class F for some neighborhood D of t_0 in M. However, S_t is injective for t near t_0 and $R(S_t) \subset N(T_t)$. Hence $\dim R(S_t) = \dim N_0 = \dim N(T_t)$ for t in D, which implies that $R(S_t) = N(T_t)$ for such t and $E_t = P_t$. Hence the function $t \to P_t \colon D \to L(H_1, H_1)$ is of class F.

However, we also know from the proof of Lemma 1(b) that $P_t = E_t = S_t X_t S_t^*$, where $t \to S_t^* \colon M \to L(H, N_0)$ is of class F, $t \to X_t \colon D \to L(N_0, N_0)$ is of class F, while $t \to S_t \colon M \to L(N_0, W)$ is of class F. Hence $t \to E_t \colon M \to L(H_1, W)$ is of class F.

LEMMA 4. *Let H_1 and H_2 be two Hilbert spaces, $\{T_t \colon t \in M\}$ a family of closed linear operators from H_1 to H_2 such that each T_t has a closed range $R(T_t)$ in H_2 while the domain of T_t is a fixed linear subset W of H_1, where W carries a Hilbert space structure with a continuous injection into H_1. Suppose that* $\dim N(T_t)$ *is finite and independent of t in M. Let Q_t be the orthogonal projection map of H_2 on $R(T_t)$.*

Then the function $t \to Q_t \colon M \to L(H_2, H_2)$ is of class F.

Proof of Lemma 4. It suffices to prove for a given t_0 in M that there exists a neighborhood D of t_0 such that the function $t \to Q_t \colon D \to L(H_2, H_2)$ is of class F. Let $R_0 = R(T_{t_0})$, $N_0 = N(T_{t_0})$, C_0 the orthogonal complement of N_0 in W, P'_t the orthogonal projection map of W on $N(T_t)$. Then T_{t_0} is an isomorphism of C_0 on R_0. Let $G_0 \in L(R_0, W)$ be the inverse of this

isomorphism, and let $S_t = T_t(I - P'_t) G_0 \colon R_0 \to H_2$. The function $t \to P'_t\colon M \to L(W, W)$ is of class F and hence the function $t \to S_t\colon M \to L(R_0, H_2)$ is of class F. Since S_{t_0} is the injection map J of R_0 into H_2, we may apply Lemma 1 and obtain the conclusion that the map $t \to E_t\colon D \to L(H_2, H_2)$ is of class F for some neighborhood D of t_0, where E_t is the orthogonal projection of H_2 on $R(S_t)$. However, the map K_t of W into W given by $K_t w = (I - P'_t)(I - P'_{t_0}) + P'_t$ is continuous from D to $L(W, W)$ with $\| K_t - I \| < 1$ for t near t_0. If $C_t = K_t(C_0)$, we see that $W = C_t + N_t$ since for t near t_0, P'_t maps N_0 isomorphically onto N_t. Hence $T_t(C_t) = R(S_t) = R(T_t)$ for t near t_0, $E_t = Q_t$, and the Lemma is proved.

LEMMA 5. *Let H_1 and H_2 be two Hilbert spaces, $\{T_t\colon t \in M\}$ a family of closed linear operators from H_1 to H_2 such that each T_t has a closed range in H_2 while the domain of T_t is a fixed linear subset W of H_1 carrying a Hilbert space structure with a continuous injection into H_1. Suppose that $\dim \mathrm{N}(T_t)$ is finite and independent of t on M. Let G_t be the generalized inverse of T_t. Suppose that the function $t \to T_t\colon M \to L(W, H_2)$ is of class F.*

Then the function $t \to G_t\colon M \to L(H_2, W)$ is of class F.

Proof of Lemma 5. We need consider G_t only for t on a small neighborhood D of a given t_0 in M. Let G_0 be the restriction of G_{0_0} to $R_0 = R(T_{t_0})$. By the proof of Lemma 4, for t near t_0 $T_t G_0 = T_t(I - P'_t) G_0 = S_t$ has the same range as T_t. Set $U_t = (I - P_t) G_0 W_t$ where $W_t\colon H_2 \to R_0$ satisfies the equation $T_t G_0 W_t = Q_t$. Then $P_t U_t u = 0$ and $T_t U_t u = T_t G_0 W_t u = Q_t u$ for each u in H_2, i. e. $U_t = G_t$.

If $S_t\colon R_0 \to H_2$ is defined by $S_t = T_t G_0$, the equation $T_t G_0 W_t = Q_t$ becomes $S_t W_t = Q_t$ and the range of Q_t coincides with the ranges of S_t for t near t_0. Moreover, the functions $t \to S_t\colon M \to L(R_0, H_2)$ and $t \to Q_t\colon M \to L(H_2, H_2)$ are both of class F. Hence we may apply Lemma 1(c) and assert that for t near t_0, there exists an unique element W_t in $L(H_2, R_0)$ such that $S_t W_t = Q_t$ and that the function $t \to W_t\colon D \to L(H_2, R_0)$ is of class F for some neighborhood D of t_0. Hence the function $t \to K_t = U_t = (I - P_t) G_0 W_t\colon D \to L(H_2, W)$ is of class F since by Lemma 3, the function $t \to (I - P_t)\colon M \to L(H_1, W)$ is of class F.

UNIVERSITY OF CHICAGO.

REFERENCES.

[1] F. E. Browder, "Analyticity and partial differential equations, I," *American Journal of Mathematics*, vol. 84 (1962), pp. 666-710.

[2] I. E. Gokhberg and M. G. Krein, "Fundamental aspects of defect numbers, root numbers, and indices for linear operators," *Uspekhi Matematicceskih Nauk SSSR* (N. S.), vol. 12 (1957), pp. 43-118 (*A. M. S. Translations*, Series No. 2, vol. 13, pp. 185-264).

[3] L. Hörmander, *Linear partial differential operators* (Die Grundlehren der Mathematischen Wissenschaften, Bd. 116), Berlin, 1963.

[4] K. Kodaira and D. C. Spencer, "On deformations of complex structures, III. Stability theorems for complex structures," *Annals of Mathematics*, vol. 71 (1960), pp. 43-76.

[5] A. Weil, "On discrete subgroups of Lie groups (II)," *Annals of Mathematics* (2), vol. 75 (1962), p. 578-602.

CORRECTION TO "FINITELY GENERATED KLEINIAN GROUPS."

By LARS AHLFORS.

Lipman Bers has pointed out to me that the principal result in my paper is too optimistically formulated. I prove, correctly, that a finitely generated group has only a finite number of linearly independent quadratic differentials, but I am not justified in concluding that the orbit space Ω/Γ has only finitely many components. My mistake was to overlook the fact that a sphere with three punctures has a Poincaré metric, but no quadratic differentials. The reasoning does not exclude the possibility of infinitely many components of this type, and it remains an open question whether this can actually occur.

The mistake recurs in the proof of Theorem 6, but the assertion remains valid. Indeed, it has been shown by R. Accola (to appear) that in the presence of two invariant regions all other components of Ω are "atoms" in the sense that they are invariant only under the identity transformation. Since atoms cannot be triply punctured spheres the error does not influence the reasoning.

On p. 418, line 13, read "restrictions" instead of "restriction," and on the same page, line 23, replace "O" by "o."

REFERENCE.

[1] Lars V. Ahlfors, "Finitely generated Kleinian Groups," *American Journal of Mathematics*, vol. 86 (1964), pp. 413-429.

Received February 10, 1965.

ERRATA.

Errata to Arthur Sard, "Hausdorff measure of critical images on Banach manifolds." *American Journal of Mathematics*, vol. 87 (1965), pp. 158-174.

Page 160, line 11. For "$A = \cdots$" read "$\theta A = \cdots$."

Page 162, second line before Theorem 1. For "$1, 2, \cdots, m$" read "$0, 1, \cdots, m$."

Page 170. Delete the Corollary and its two following lines, as all appear on page 169 in their proper place.

Page 173, line following Statement 3. For "$r = m - 1$, $r = n - 1$," read "$r \geqq m - 1$, $r \geqq n - 1$."

Page 166, Part 2. For "Tynchonoff" read "Tychonoff."

Page 174, [20]. For "*La théorie de . . .*" read "*La théorie des . . .*".

Page 174, [21]. For "Siracura" read "Siracusa."

Received June 21, 1965.

ON t-DESIGNS AND GROUPS.

By D. R. HUGHES.[1]

1. Introduction. As Ryser has pointed out in his recent book, combinatorial analysis can be thought of as the study of how to put objects together into sets satisfying certain restrictions ([9]). A type of restriction that has been heavily studied is that of insisting that any two of the objects lie in the same number of common subsets; a much stronger restriction seems to be the demand that any t of the objects lie in the same number of common subsets, where $t \geqq 2$. As we shall see, this is a problem closely related to the problem of transitive extensions of groups, but the appeal is still strongly combinatorial. "Designs" of the kind specified will be called *t-designs* if they also have the porperty that any two of the subsets have the same number of elements. No non-trivial 6-designs are known, and only a finite number of 5-designs and 4-designs have been found (they will all be found here, or referred to at least). If it could be shown that there are only trivial t-designs, for some sufficiently large t, it would imply that the only (finite) t-fold transitive permutation groups are alternating or symmetric.

In Section 2 we give some definitions and simple basic facts about t-designs. In Section 3 we introduce the group theory that we will be concerned with, and prove several theorems about the connections between permutation groups and t-designs. Essentially, these are techniques for constructing t-designs from groups, and more especially, for having some control over the kind of design constructed (for as we shall see, it is easy to construct the designs from the groups—the problem is to know something about the parameters of the design). But these theorems also have some purely group theory applications, giving information about the possibility of transitively extending a group. In Section 4 we apply these methods to the linear fractional groups, and recall the application of these methods to the same groups in [8]. Finally, in Section 5 we consider the Mathieu groups, in more than one way, and say a little about the Suzuki groups.

Received April 13, 1964.

[1] Supported in part by the National Science Foundation under grant NSF G-18912. The author would like to express his gratitude for the support in carrying out this research given by the Department of Scientific and Industrial Research of the United Kingdom.

Besides the material on the Mathieu groups, which was first given in slightly different form by Witt ([12]), the material on the Suzuki groups has been treated before, in [6]; and as pointed out above, [8] contains some additional treatment of the linear fractional groups, though most of it is at least summarized here. The connection with inversive planes, pointed out in [6], was also noted by Tits in [11], but in a completely different way, and has been involved in several developments in the theory of inversive planes, of which we can only refer the reader to [3], for a fairly complete bibliography as of this moment. Throughout the paper permutations are to act on the right, and a representation on (right) cosets Hx will always be intended when such references are appropriate.

2. Definitions and basic results. By a *t-design* π we mean a finite set π of *points* and *blocks*, with an *incidence relation* between points and blocks (for which we always use ordinary geometric terminology such as point on block, etc.) such that there are integers λ, t, k, for which:

(i) each block contains exactly k points

(ii) each set of t distinct points is on exactly λ common blocks

(iii) if two blocks consist of exactly the same points, then they are equal as blocks.

If the number of points in π is v, then we shall say that π is a t-design for (v, k, λ), or is a t-(v, k, λ). We will use b as a symbol for the number of blocks, and point out that in [6] we wrote, for instance, t-(v, b, k, λ) where now we omit the b. For $\lambda = 1$, t-designs are just the Steiner systems of Witt ([13]), while the 2-designs are exactly the balanced incomplete block designs, symmetric if and only if $v = b$. We shall say that a t-design for (v, k, λ) is *trivial* if every subset of the appropriate size is in fact a block; i.e., the number of blocks is $b = C_{v,k}$, the binomial coefficient. A *collineation* of a t-design is a one-to-one mapping of points onto points, blocks onto blocks, which preserves incidence. A collineation group of π will be called s-transitive if it is s-fold transitive when considered as a permutation group on the points. The *flags* of π are the couples (P, a), where P is a point and a is a block containing P. A group is *flag-transitive* if it is transitive on the flags; this is equivalent to the assertion that the group is transitive on blocks (points) and the subgroup fixing a block (point) is transitive on the points of that block (the blocks through that point). A collineation group is *s-flag-transitive* if it is transitive on blocks and the subgroup fixing a block is s-fold transitive on the points of the block.

There are many elementary counting theorems about t-designs, and we give several here. Some of these are to be found in [6], and also compare [2] for some similar results.

LEMMA 2.1. *If π is a t-(v,k,λ), and s is an integer, $0<s<t$, then every set of s distinct points of π is on exactly λ_s common blocks, where*

$$\lambda_s=\lambda(v-s)(v-s-1)\cdots(v-t+1)/(k-s)(k-s-1)\cdots(k-t+1),$$

and so π is also an s-(v,k,λ_s).

Proof. A set S of s distinct points can be completed to a set of t distinct points in exactly $C_{v-s,t-s}$ ways, each way giving us λ blocks on S. But each block on S contains $C_{k-s,t-s}$ subsets of $t-s$ points which define the same set of blocks, and so the number of blocks on S is $\lambda C_{v-s,t-s}/C_{k-s,t-s}$, and upon simplification, this gives the formula of the lemma.

LEMMA 2.2. *If π is a t-(v,k,λ), with b blocks, then $bk=v\lambda_1$.*

Proof. The right side of $bk=v\lambda_1$ counts the number of points, times the number of blocks per point, while the left side counts the number of blocks, times the number of points per block; that is, both sides count the number of "incidences" in π.

Now if we understand that $\lambda_t=\lambda$, $\lambda_0=b$, then the formula of Lemma 2.1 is valid for $0\leqq s\leqq t$.

LEMMA 2.3. *If $t\geqq 2$, then $b\geqq v$.*

Proof. This is equivalent to the assertion that a 2-design has at least as many blocks as points, and this well-known result can be found, for instance, in [9].

Given a t-design π for (v,k,λ), and a point P in π, we define π_P to be the object consisting of those points of π other than P, and of those blocks of π which contain P. Using the natural incidence relation, it is clear that π_P is a $(t-1)$-design for $(v-1,k-1,\lambda)$, with λ_1 blocks. We say that a t-design π is an *extension* of a $(t-1)$-design π', if π' is isomorphic to π_P, for some point P in π; thus π might be an extension of many different (i.e., non-isomorphic) $(t-1)$-designs. Conversely, π_P is a *restriction* of π. By considering $t-2$ successive restrictions, Lemma 2.3 immediately permits:

LEMMA 2.4. *If $t\geqq i+2$, then $\lambda_i\geqq v-i$.*

Now combining Lemmas 2.1 and 2.4 we have:

LEMMA 2.5. $bk(k-1)\cdots(k-t+3)\geqq v(v-1)\cdots(v-t+2)$.

It is well-known that if $t=2$, then there are infinitely many examples of designs for which $v=b$; but:

LEMMA 2.6. (See also [2], Satz 5). *If $b=v$ and $t>2$, then $k\geqq v-1$ and π is trivial.*

Proof. Suppose π is a t-(v,k,λ), with $b=v$. If $t\geqq 3$, then from Lemma 2.4, $\lambda_1\geqq v-1$. But $\lambda_1=bk/v=vk/v=k$. If $k=v$, then there is only one block and certainly π is trivial (in fact, then $b=v=1$). If $k=v-1$, then $C_{b,k}=C_{v,v-1}=v=b$, so π is trivial.

A t-design π, where $t\geqq 2$, will be called *symmetric* if b is as small as possible in the light of Lemma 2.5. That is, a symmetric t-design, for $t>2$, is one which is an extension of a symmetric $(t-1)$-design, and so $bk(k-1)\cdots(k-t+3)=v(v-1)\cdots(v-t+2)$; a symmetric 2-design is one for which $b=v$.

THEOREM 2.1. *For a fixed value of λ and a fixed $t>2$, there are only finitely many symmetric t-designs for (v,k,λ).*

Proof. Since π is symmetric,

$$b=v(v-1)\cdots(v-t+2)/k(k-1)\cdots(k-t+3). \tag{1}$$

Also, since (1) above and Lemma 2.1 and 2.2 give us two different expressions for b, we set them equal and obtain,

$$\lambda(v-t+1)=(k-t+2)(k-t+1). \tag{2}$$

Hence from (2) $v=t-1+A/\lambda$, where $A=(k-t+1)(k-t+2)$, and so,

$$v(v-1)\cdots(v-t+2)=(A+\lambda)(A+2\lambda)\cdots(A+(t-1)\lambda)/\lambda^{t-1}, \tag{3}$$

But k divides the left hand side of (3), from (1), and so k divides the numerator of the right hand side. Since $A\equiv(t-1)(t-2)\pmod{k}$, this yields

$$\begin{aligned}&[(t-1)(t-2)+\lambda][(t-1)(t-2)+2\lambda]\cdots\\ &\qquad\times[(t-1)(t-2)+(t-1)\lambda]\equiv 0 \pmod{k}.\end{aligned} \tag{4}$$

Then k is bounded by the left hand side of (4); hence, from (2), v is bounded when t and λ are specified (if $t>2$; note that (1) becomes $b=v$ if $t=2$, and the argument above fails, as of course it must).

(*Remark.* The author would like to thank Peter Dembowski for pointing out an error in both the statement and proof of an earlier form of the theorem above.)

3. Some group theory. We first explain some general terminology, which will be modified or elaborated in particular situations. All groups considered here are finite, and we use $|\mathfrak{H}|$ and $|\mathfrak{H}:\mathfrak{K}|$ to denote, respectively, the order of the group $\mathfrak{H}$ and the index in $\mathfrak{H}$ of the subgroup $\mathfrak{K}$; in general we use $|\mathcal{S}|$ to denote the number of elements in the set $\mathcal{S}$. Following the suggestion made in the previous section, we will say that a t-fold transitive permutation group is, briefly, t-transitive; frequently we will denote such a group by $\mathfrak{G}^t$. Given a permutation group $\mathfrak{G}$, if P is one of the symbols on which $\mathfrak{G}$ acts, then $\mathfrak{G}_P$ is the subgroup of $\mathfrak{G}$ fixing P, and we shall not consider P, in general, as one of the symbols on which $\mathfrak{G}_P$ acts, so that we can speak of $\mathfrak{G}_P$ being transitive, for instance. When $\mathfrak{G}$ is t-transitive, then of course $\mathfrak{G}_P$ is $(t-1)$-transitive, and if we denote $\mathfrak{G}$ by $\mathfrak{G}^t$ we shall denote $\mathfrak{G}_P$ by $\mathfrak{G}^{t-1}$; this is of course not a completely unambiguous notation, but its use is usually clear. We will, for instance, sometimes consider the subgroup chain $\mathfrak{G}^t \supset \mathfrak{G}^{t-1} \supset \cdots \supset \mathfrak{G}^i$, where it will be understood that we have chosen some sequence $P_t, P_{t-1}, \cdots, P_{i+1}$ of symbols to define the respective subgroups.

Now if π is a t-design, $\mathfrak{G}$ a collineation group of π, and P a point of π, then it is clear that $\mathfrak{G}_P$ is a collineation group of π_P. If $\mathfrak{G}$ is s-transitive on π, then obviously $\mathfrak{G}_P$ is $(s-1)$-transitive on π_P, and it is a sort of converse of this that we wish to study for our first theorem. Recall that a group $\mathfrak{A}$ is called a *transitive extension* of a group $\mathfrak{B}$ if $\mathfrak{A}$ is a transitive permutation group and the subgroup fixing a symbol is permutation-isomorphic to $\mathfrak{B}$.

Suppose $\mathfrak{G}^t$ is a t-transitive collineation group of the t-design π for (v,k,λ), and that also $\mathfrak{G}^t$ is transitive on the blocks of π. Suppose also that there is a group $\mathfrak{G}^{t+1}$ which is a transitive extension of $\mathfrak{G}^t$, thinking of the latter as a group on the points; that is, $\mathfrak{G}^{t+1}$ acts on the v points of π, plus a new symbol Q. Let d be some fixed block of π, and $\mathfrak{J}^t$ the subgroup of $\mathfrak{G}^t$ which fixes d; let d' be the set consisting of the points of d, plus the new symbol Q. Let $\mathfrak{J}^{t+1}$ be the subgroup of $\mathfrak{G}^{t+1}$ which fixes the set d' (as a set of course). Now we construct a $(t+1)$-design π' by using for points the points of π, plus Q, and for blocks the images of d', under $\mathfrak{G}^{t+1}$.

THEOREM 3.1. *π' is a $(t+1)$-design for $(v+1, k+1, \lambda')$, where $\lambda' = \lambda(k+1)/|\mathfrak{J}^{t+1}:\mathfrak{J}^t|$, and $\mathfrak{G}^{t+1}$ is a $(t+1)$-transitive collineation group of π', and is transitive on the blocks of π'.*

COROLLARY 1. *If π has b blocks and π' has b' blocks, then*

$$b' = b(v+1)/|\mathfrak{J}^{t+1}:\mathfrak{J}^t|.$$

COROLLARY 2. *$(\pi')_Q$ is isomorphic to π if and only if $\mathfrak{G}^{t+1}$ is flag-transitive, or equivalently,* $|\mathfrak{J}^{t+1}:\mathfrak{J}^t| = k+1$.

COROLLARY 3. *If $\mathfrak{G}^t$ is flag-transitive, then either $\lambda' = \lambda$, or $\lambda' = \lambda(k+1)$*

Proofs. The theorem and first corollary were proved in [6], and we do not repeat the proofs here. Corollary 2 follows immediately from observing that $(\pi')_Q$ is isomorphic to π if and only if $\lambda' = \lambda$, and this means $\mathfrak{J}^t$ must have index $k+1$ in $\mathfrak{J}^{t+1}$; but this last is exactly the statement that $\mathfrak{J}^{t+1}$ is transitive on the points of the block d', since $\mathfrak{J}^t$ is the subgroup fixing one point, Q. Hence $\mathfrak{G}^{t+1}$ must be flag-transitive, and conversely. Corollary 3 follows from the observation that if $\mathfrak{J}^t$ is transitive on the points of d, then the orbit of Q, under $\mathfrak{J}^{t+1}$, either consists of Q alone or of all of d'.

It is in fact straight-forward to compute that if λ'_i means the number of blocks on i points in π', then

$$\lambda'_i = \lambda_i(v+1-i)(k+1)/(k+1-i)\,|\mathfrak{J}^{t+1}:\mathfrak{J}^t|.$$

When there is some possibility of computing $|\mathfrak{J}^{t+1}:\mathfrak{J}^t|$, Theorem 3.1 gives us a way to construct t-designs from t-transitive groups. We start at some convenient point, for instance we make the subgroup fixing $t-1$ symbols define a 1-design in some fashion, and then repeatedly apply Theorem 3.1, finally arriving at a t-design. Now in fact, it is apparent that if we are given a t-transitive group and choose any subset (with at least t points) whatsoever, and use this subset and its images as blocks, we will have a t-design. Will the design constructed in this fashion be trivial? If so then the group is transitive on the sets of the appropriate size; i.e., is *k-homogeneous,* where k is the size of our proposed blocks. But if the group is k-homogeneous for all k equal to or less than its degree, then the results of Beaumont and Peterson ([1]) assure us that the group is alternating or symmetric, or is one of four certain groups, all of degree less than 10, and each no more than 3-transitive. Furthermore, the results of [7] assure us that if k is small enough compared to the degree, then the group cannot be k-homogeneous without being $(k-1)$-transitive. So, excepting the four groups of [1], mentioned above, any t-transitive group is associated with at least one non-trivial t-design, unless the group is alternating or symmetric. But the problem of finding the parameters is more difficult.

Suppose π is a t-design for (v,k,λ), and $\mathfrak{G}^t$ is a t-transitive group on π, $\mathfrak{G}^t$ also transitive on blocks. Let P be a point of π, d a block of π, let $\mathfrak{G}^{t-1} = (\mathfrak{G}^t)_P$, and $\mathfrak{J}^t$ the subgroup of $\mathfrak{G}^t$ fixing the block d. Define $\mathfrak{D}$ to be the set of all elements of $\mathfrak{G}^t$ which send P onto d. Then clearly $\mathfrak{D} = \mathfrak{G}^{t-1}\mathfrak{D}\mathfrak{J}^t$,

and so $\mathfrak{D}$ consists of k right cosets of $\mathfrak{G}^{t-1}$, and also consists of λ_1 left cosets of $\mathfrak{J}^t$.

Conversely, given $\mathfrak{G}^t$, $\mathfrak{G}^{t-1}$, and $\mathfrak{D}$, we can reconstruct π, for d will consist of those right cosets of $\mathfrak{G}^{t-1}$ which are in $\mathfrak{D}$, and $\mathfrak{J}^t$ will be the subgroup of all x such that $\mathfrak{D}x = \mathfrak{D}$; i.e., $\mathfrak{J}^t$ is the subgroup fixing $\mathfrak{D}$. It is clear that $\mathfrak{D}$ can be any union of right cosets of $\mathfrak{G}^{t-1}$, since any such choice simply corresponds to a choice of a subset to be a block, and then the demand that $\mathfrak{G}^t$ be transitive on blocks.

LEMMA 3.1. *Let π be the t-design whose points are the right cosets of $\mathfrak{G}^{t-1}$ and whose blocks are the images of $\mathfrak{D}$, where $\mathfrak{D}$ is a union of right cosets of $\mathfrak{G}^{t-1}$. Then if $\mathfrak{J}^t$ is the subgroup fixing $\mathfrak{D}$ and if we assume (obviously without loss of generality) that $\mathfrak{G}^{t-1}$ itself is one of the right cosets in $\mathfrak{D}$, it follows that $\mathfrak{G}^{t-1}\mathfrak{J}^t \subset \mathfrak{D}$. $\mathfrak{G}^t$ is flag-transitive on π if and only if $\mathfrak{G}^{t-1}\mathfrak{J}^t = \mathfrak{D}$.*

Proof. Let k be the number of right cosets of $\mathfrak{G}^{t-1}$ in $\mathfrak{D}$; then $\mathfrak{G}^t$ is flag-transitive if and only if the index in $\mathfrak{J}^t$ of the subgroup fixing a point on the block $\mathfrak{D}$ is k. But the subgroup fixing a point can be taken to be $\mathfrak{G}^{t-1} \cap \mathfrak{J}^t$, and since $|\mathfrak{J}^t : \mathfrak{G}^{t-1} \cap \mathfrak{J}^t| = |\mathfrak{G}^{t-1}\mathfrak{J}^t : \mathfrak{G}^{t-1}|$ is a well-known group theory relationship, it follows that $\mathfrak{J}^t$ is transitive on the k points if and only if $\mathfrak{G}^{t-1}\mathfrak{J}^t$ consists of k right cosets of $\mathfrak{G}^{t-1}$, or if and only if $\mathfrak{G}^{t-1}\mathfrak{J}^t$ is all of $\mathfrak{D}$.

THEOREM 3.2. *Let $\mathfrak{G}^t$, $\mathfrak{G}^{t-1}$, $\mathfrak{D}$, $\mathfrak{J}^t$ and π have the same meaning as in Lemma 3.1. Then $\mathfrak{G}^t$ is flag-transitive if and only if there is a subgroup $\mathfrak{A}$ of $\mathfrak{G}^t$ such that $\mathfrak{D} = \mathfrak{G}^{t-1}\mathfrak{A}$.*

Proof. If $\mathfrak{G}^t$ is flag-transitive, then $\mathfrak{J}^t$ itself is the subgroup $\mathfrak{A}$. Conversely, suppose $\mathfrak{D} = \mathfrak{G}^{t-1}\mathfrak{A}$ for some subgroup $\mathfrak{A}$. Then since $\mathfrak{D}a = \mathfrak{D}$ for any a in $\mathfrak{A}$, certainly $\mathfrak{A} \subset \mathfrak{J}^t$. But then $\mathfrak{G}^{t-1}\mathfrak{J}^t \supset \mathfrak{G}^{t-1}\mathfrak{A} = \mathfrak{D} \supset \mathfrak{G}^{t-1}\mathfrak{J}^t$, and so $\mathfrak{D} = \mathfrak{G}^{t-1}\mathfrak{J}^t$ and by Lemma 3.1, we are done.

Thus we can even construct t-designs on which $\mathfrak{G}^t$ acts flag-transitively, by the simple expedient of choosing $\mathfrak{D} = \mathfrak{G}^{t-1}\mathfrak{A}$. This raises the interesting question of whether it is always possible to construct a non-trivial design on which the group acts flag-transitively, but we do not go further into that problem here.

THEOREM 3.3. *Let $\mathfrak{G}^t$ be a t-transitive collineation group of the t-design π for (v,k,λ). Then $\mathfrak{G}^t$ is s-flag-transitive if and only if $|\mathfrak{G}^{t-s} : \mathfrak{J}^{t-s}| = \lambda_s$. Here $\mathfrak{G}^i$ is the subgroup of $\mathfrak{G}^t$ fixing $t-i$ points, and $\mathfrak{J}^i$ is the intersection of $\mathfrak{G}^i$ with the subgroup $\mathfrak{J}^t$ fixing some prescribed block.*

Proof. If $\mathfrak{G}^t$ is s-flag-transitive, then $\mathfrak{G}^{t-s}$ is easily seen to be transitive on blocks, and so $|\mathfrak{G}^{t-s}:\mathfrak{J}^{t-s}| = \lambda_s$. For the other half, we give a proof by induction. Consider the subgroup lattice below:

$$\begin{tikzcd}[column sep=large, row sep=large]
 & \mathfrak{G}^{t-s+1} \arrow[dl, dash, "v-s+1"'] \arrow[dr, dash, "y"] & \\
\mathfrak{G}^{t-s} \arrow[dr, dash, "\lambda_s"'] & & \mathfrak{J}^{t-s+1} \arrow[dl, dash, "x"] \\
 & \mathfrak{J}^{t-s} &
\end{tikzcd}$$

Now $x \leqq k-s+1$, since $\mathfrak{J}^{t-s+1}$ is a permutation group on the $k-s+1$ points of a block, and $\mathfrak{J}^{t-s}$ is the subgroup fixing one of these points. Similarly, $y \leqq \lambda_{s-1}$, since y is the index in $\mathfrak{G}^{t-s+1}$ of the subgroup fixing one of the λ_{s-1} blocks on which $\mathfrak{G}^{t-s+1}$ acts. From the lattice, $xy = \lambda_s(v-s+1)$. But from Lemma 2.1, we have $\lambda_s(v-s+1) = \lambda_{s-1}(k-s+1)$, and so $x = k-s+1$ and $y = \lambda_{s-1}$. But then, by induction, $\mathfrak{G}^t$ is $(s-1)$-flag-transitive, or $\mathfrak{J}^t$ is $(s-1)$-transitive on the points of a block. But $\mathfrak{J}^{t-s+1}$ is the subgroup of $\mathfrak{J}^t$ fixing $s-1$ points, and since $x = k-s+1$, it follows that $\mathfrak{J}^{t-s+1}$ is transitive, so $\mathfrak{J}^t$ is s-transitive and $\mathfrak{G}^t$ is s-flag-transitive. Note that if $s=1$ this proof still works, and simply shows that $\mathfrak{G}^t$ is transitive on blocks as well as on the points of a block.

The next theorem is a generalization of a method used by Witt in ([12]) to analyze some of the Mathieu groups. If A is a group, B a subgroup of A, and if, for all (finite) groups H which contain A, it is true that whenever $h^{-1}Bh$ is in A for some h in H, then $h^{-1}Bh = x^{-1}Bx$ for some x in A, then we call B an *S-subgroup* of A. A itself, as well as all of its Sylow subgroups, are S-subgroups, for example. A well-known group theory result (see [4], page 68) can trivially be generalized to:

THEOREM I. *If $\mathfrak{G}^t$ is a t-transitive group, $\mathfrak{G}^0$ the subgroup fixing t letters, and $\mathfrak{B}$ an S-subgroup of $\mathfrak{G}^0$ such that $\mathfrak{B}$ fixes $k \geqq t$ letters, then the normalizer of $\mathfrak{B}$ in $\mathfrak{G}^t$ is t-transitive on the k letters fixed.*

THEOREM 3.4. *Let $\mathfrak{G}^t$ be a t-transitive permutation group on v letters and $\mathfrak{B}$ an S-subgroup of $\mathfrak{G}^0$ such that $\mathfrak{B}$ fixes $k > t$ letters; let $\mathcal{K}$ be this set of k letters. Then if $\mathcal{K}$ and its images under $\mathfrak{G}^t$ are the blocks, we have*

a t-design for (v,k,λ), *where* λ *is the index in* $\mathfrak{G}^0$ *of the subgroup of* $\mathfrak{G}^0$ *which fixes* $\mathscr{K}$, *as a set.* $\mathfrak{G}^t$ *is t-flag-transitive on this design.*[2]

Proof. If $\mathfrak{J}^t$ is the subgroup of $\mathfrak{G}^t$ fixing $\mathscr{K}$, then since $\mathfrak{J}^t$ contains the normalizer of $\mathfrak{B}$ in $\mathfrak{G}^t$, and this normalizer, by Theorem I, is t-transitive on $\mathscr{K}$, certainly $\mathfrak{J}^t$ is; hence $\mathfrak{G}^t$ is t-flag-transitive. So, from Theorem 3.3, $\lambda=\lambda_t=|\mathfrak{G}^0:\mathfrak{J}^0|$, where $\mathfrak{J}^0=\mathfrak{G}^0\cap\mathfrak{J}^t$; clearly $\mathfrak{J}^0$ is the subgroup of $\mathfrak{G}^0$ fixing $\mathscr{K}$ as a set.

COROLLARY 1. (Witt, [13]) .*In the situation of Theorem* 3.4, *if* $\mathfrak{B}$ *is normal in* $\mathfrak{G}^0$, *then* $\lambda=1$.

Proof. Certainly $\mathfrak{J}^0$ contains the normalizer of $\mathfrak{B}$ in $\mathfrak{G}^0$, in any case, so the corollary is immediate.

We return to a particular aspect of Theorem 3.1 which the author first noted when studying the Mathieu groups. If in Theorem 3.1 we know that λ remains constant in the extension from π to π', then it follows that $b'=(v+1)b/(k+1)$, and this restriction limits the possibility of extension. The only result we have in this direction is partial, but just because it works so well for the Mathieu groups, and indeed replaces some of the work in Zassenhaus' paper [14], it is worth giving here. But first we need some more definitions. If $\mathfrak{A}$ is a group and $\mathfrak{B}$, $\mathfrak{C}$ are subgroups of $\mathfrak{A}$, then we say that $\mathfrak{C}$ is $\mathfrak{B}$-*special* if every automorphism of $\mathfrak{A}$ which fixes $\mathfrak{B}$ also fixes $\mathfrak{C}$. For example, if $\mathfrak{C}$ is the product of $\mathfrak{B}$ and a characteristic subgroup of $\mathfrak{A}$, then $\mathfrak{C}$ is $\mathfrak{B}$-special.

In [12], Witt has given us a precise theory of transitive extensions, and while the whole theory is interesting, we only sketch the parts that are essential to our purpose here. Given a t-transitive group $\mathfrak{G}^t$ on letters $P_1,P_2,\cdots,P_t,Q_1,Q_2,\cdots$, we choose the $\mathfrak{G}^i$ so that $\mathfrak{G}^i$ is the subgroup of $\mathfrak{G}^{i+1}$ fixing P_{i+1}, for $i=0,1,\cdots,t-1$. Then one can find elements $\sigma_i=(P_i,P_{i-1})\rho_i$, where ρ_i is a permutation moving only the symbols Q_j (and in fact, in the finite case, ρ_i can be chosen to have order 2), for $i=2,\cdots,t$, such that

(1) $\mathfrak{G}^i=\mathfrak{G}^{i-1}+\mathfrak{G}^{i-1}\sigma_i\mathfrak{G}^{i-1}$, for $i=2,\cdots,t$

(2) $\sigma_i\mathfrak{G}^1\sigma_i=\sigma_i^{-1}\mathfrak{G}^1\sigma_i=\mathfrak{G}^1$, if $i\geqq 3$

(3) $(\sigma_i\sigma_j)^{m_{ij}}\in\mathfrak{G}^0$, where $m_{ij}=1,2,$ or 3, according as $i=j$, $|i-j|>1$, or $|i-j|=1$.

Here, in (1), we use $+$ to denote set union.

[2] The theorem is actually still valid if $k=t$, but the design is trivial, as must always be the case for a t-design with $k=t$.

Conversely, if we have a 2-transitive group $\mathfrak{G}^2$ acting on the symbols $P_1, P_2, Q_1, Q_2, \cdots$, and if $\mathfrak{G}^1 = (\mathfrak{G}^2)_{P_2}$, $\mathfrak{G}^0 = (\mathfrak{G}^1)_{P_1}$, $\sigma_2 = (P_2\ P_1)\rho_2$, where ρ_2 moves only the Q_j, where σ_2 is in $\mathfrak{G}^2$, and if we have elements σ_i as above, satisfying (2) and (3), and if we recursively define the $\mathfrak{G}^i$ by (1), then each $\mathfrak{G}^i$ is i-transitive and is a transitive extension of $\mathfrak{G}^{i-1}$, for $i = 3, \cdots, t$.

THEOREM 3.5. *Suppose $\mathfrak{G}^2$ is a 2-transitive group on the 2-design π^2 for $(v, k, 1)$, and that* (i) *$\mathfrak{G}^2$ is 2-flag-transitive, and* (ii) *the subgroup $\mathfrak{J}^1$ of $\mathfrak{G}^1$ fixing a block is a $\mathfrak{G}^0$-special subgroup of $\mathfrak{G}^1$. Then if $\mathfrak{G}^2$ can be successively extended to 3-, 4-, $\cdots$, t-transitive groups $\mathfrak{G}^3, \mathfrak{G}^4, \cdots, \mathfrak{G}^t$, the 3-, 4-, $\cdots$, t-designs $\pi^3, \pi^4, \cdots, \pi^t$, defined in this fashion by Theorem* 3.1 *will all have $\lambda = 1$.*

Proof. Under the hypotheses of the theorem, we have elements σ_i satisfying (1), (2), and (3). But it is easy to see that for $i = 2$, the σ_i also have the property $\sigma_i^{-1}\mathfrak{G}^0\sigma_i = \mathfrak{G}^0$, and since $\mathfrak{J}^1$ is $\mathfrak{G}^0$-special in $\mathfrak{G}^1$, the automorphisms σ_i of $\mathfrak{G}^1$ must fix $\mathfrak{J}^1$ as well. Since $\mathfrak{J}^0 = \mathfrak{G}^0$, they fix $\mathfrak{J}^0$ as well (this is because $\lambda = 1$). Now since $\mathfrak{J}^2$ is 2-transitive, the element σ_2 can be chosen in $\mathfrak{J}^2$, and now if we write $\mathfrak{J}^i$ instead of $\mathfrak{G}^i$ at each point in (2) and (3), we see that all the conditions are satisfied for $\mathfrak{J}^2$ to have successive transitive extensions up to a t-transitive group $\mathfrak{J}^t$.

Consider the extension from $\mathfrak{G}^i$ to $\mathfrak{G}^{i+1}$. If, inductively, we assume that $\lambda = 1$ in π^i, then $\mathfrak{G}^i$ is flag-transitive, and since $\mathfrak{J}^{i+1}$ is strictly greater than $\mathfrak{J}^i$, it follows from Corollary 3 to Theorem 3.1 that $\lambda = 1$ in π^{i+1}.

COROLLARY 1. *Under the hypotheses of Theorem* 3.5, *$\mathfrak{G}^t$ is t-flag-transitive on π^t.*

COROLLARY 2. *Under the hypotheses of Theorem* 3.5, *each of the quantities*

$$(v-1)/(k-1), v(v-1)/k(k-1), \cdots,$$
$$(v+t-2)\cdots(v-1)/(k+t-2)\cdots(k-1)$$

is an integer.

Proof. The first corollary is of course obvious. The second is merely the statement that the formula of Corollary 1 to Theorem 3.1 gives us the number of blocks in the various π^i.

4. Linear fractional groups. Here we will utilize the well-known linear-fractional groups (over finite fields) to construct designs; since the groups are 3-transitive, we will in general find 3-designs. The first approach

will be very simple and straight-forward, utilizing Theorem 3.1 in the most obvious manner.

Let q be a prime-power, $\mathfrak{F}$ the field $GF(q)$, and $\mathfrak{F}^*$ the set of non-zero elements of $\mathfrak{F}$. Then $\mathfrak{G}^1$ is a transitive group on the $q-1$ marks of $\mathfrak{F}^*$, given by the mappings $x \to ax$, where $a \neq 0$. $\mathfrak{G}^2$ is a transitive extension of $\mathfrak{G}^1$, on the q marks of $\mathfrak{F}$, given by the mappings $x \to ax+b$, $a \neq 0$. $\mathfrak{G}^3$ is a transitive extension of $\mathfrak{G}^2$, on the q marks of $\mathfrak{F}$, plus the new symbol ∞, given by the mappings $x \to (ax+b)/(cx+d)$, where $ad-bc \neq 0$.

Let $\mathfrak{J}^1$, of order $k-1$, be a multiplicative subgroup of $\mathfrak{F}^*$; we will construct a 1-design for $(q-1, k-1, 1)$, by choosing a block m_1 to consist of the marks of $\mathfrak{J}^1$, and letting the other blocks be the marks in the various cosets of $\mathfrak{J}^1$ under $\mathfrak{G}^1$. This design, π^1, has $(q-1)/(k-1)$ blocks and of course $\mathfrak{G}^1$ is flag-transitive.

Then $\mathfrak{G}^2$ defines a 2-design for (q, k, λ'), and we need to evaluate $|\mathfrak{J}^2 : \mathfrak{J}^1|$, where $\mathfrak{J}^2$ is the subgroup of $\mathfrak{G}^2$ fixing a block m_2. We choose m_2 so that it consists of m_1 plus the new symbol 0, and we represent the mapping $x \to ax+b$ by $\phi_{a,b}$. Then $\phi_{a,b}$ is in $\mathfrak{J}^2$ if and only if $ha+b$ is in m_2 for all h in m_2. Clearly $\phi_{a,0}$ is in $\mathfrak{J}^2$ if and only if a is in $\mathfrak{J}^1$, so let us examine the possibility that there is an element $\phi_{a,b}$ in $\mathfrak{J}^2$, with $b \neq 0$. Then $\mathfrak{J}^2$ is transitive on m_2 (for 0 can be moved, and $\mathfrak{J}^1$, the subgroup fixing 0, is transitive on the rest of the symbols in m_2); since the subgroup $\mathfrak{J}^1$ fixing a point is also transitive and even regular, $\mathfrak{J}^2$ is a 2-transitive Frobenius group. Now the only non-trivial Frobenius subgroups of a Frobenius group must occur in the "natural" way, which is to say that $\mathfrak{J}^2$ must be the Frobenius subgroup of $\mathfrak{G}^2$ arising from a subfield. So k must be a prime-power, $k = p$, and q is a power of k; clearly in fact this numerical condition is necessary and sufficient for $\mathfrak{J}^2$ to be larger than $\mathfrak{J}^1$. (A direct proof can be given, without using any concepts from group theory, by the way.)

LEMMA 4.1. *π^2 is a 2-design for $(p^n, p, 1)$, with $p^{n-1}(p^n-1)/(p-1)$ blocks, where $q = p^n$, p a prime-power, if and only if $\mathfrak{J}^1$ is the multiplicative group of a subfield $GF(p)$ of $\mathfrak{F}$, and in this case $\mathfrak{G}^2$ is 2-flag-transitive on π^2. If $\mathfrak{J}^1$ is not such a subgroup, then π^2 is a 2-design for (q, k, k), with $q(q-1)/(k-1)$ blocks.*

Next we examine π^3, the design coming from $\mathfrak{G}^3$. We let m_3 be the block obtained by adjoining ∞ to m_2, and $\mathfrak{J}^3$ the subgroup fixing it. As before, we must determine $|\mathfrak{J}^3 : \mathfrak{J}^2|$. First we note that the mapping $x \to 1/x$ is always in $\mathfrak{J}^3$, and so in all cases, $\mathfrak{J}^3 \neq \mathfrak{J}^2$. Now we distinguish two cases.

Case 1. If π^2 is a 2-design for $(p^n, p, 1)$, then Corollary 3 to Theorem 3.1 assures us that $|\mathfrak{J}^3 : \mathfrak{J}^2| = p + 1$, so π^3 is a 3-design for $(p^n + 1, p + 1, 1)$, with $p^{n-1}(p^{2n} - 1)/(p^2 - 1)$ blocks.

Case 2. If π^2 is a 2-design for (q, k, k), then the action of $\mathfrak{J}^3$ on m_3 can be in only one of two ways. Either $\mathfrak{J}^3$ has two orbits, one of length 2 and the other of length $k - 1$, or else $\mathfrak{J}^3$ is transitive on m_3. In the latter case, $\mathfrak{J}^3$ is a transitive permutation group on $k + 1$ letters, such that the subgroup fixing one letter in fact fixes two letters and is transitive and regular on the remaining $k - 1$ letters; furthermore the subgroup fixing one letter is abelian. We abstract from this situation a bit:

THEOREM 4.1. *Let $\mathfrak{H}^2$ be a transitive permutation group on $n > 2$ letters such that* (i) *the subgroup fixing one letter in fact fixes two letters and is transitive and regular on the remaining $n - 2$, and* (ii) *in the subgroup fixing a letter, all subgroups of order two are normal. Then $n = 4$ or 6.*

Proof. Let $\mathfrak{H}^1$ be the subgroup fixing a letter. Each conjugate of $\mathfrak{H}^1$ fixes a letter and hence two letters. These sets of two letters cannot "overlap," for if they did, there would be non-identity elements fixing three letters, which violates the hypotheses. So the sets of two letters fixed by conjugates of $\mathfrak{H}^1$ are systems of imprimitivity for $\mathfrak{H}^2$. On these systems, $\mathfrak{H}^2$ induces a 2-transitive group, and so the order of $\mathfrak{H}^2$ is $(n/2)(n/2 - 1)k$, where k is the order of the subgroup fixing two systems; call this subgroup $\mathfrak{T}_0$. Also since $\mathfrak{H}^2$ clearly has order $n(n - 2)$, we have $k = 4$.

Let $\mathfrak{N}$ be the subgroup of $\mathfrak{H}^2$ consisting of all elements which fix every system; in other words, $\mathfrak{N}$ is the kernel of the representation of $\mathfrak{H}^2$ on the systems, and of course, $\mathfrak{N} \subset \mathfrak{T}_0$. Let $\mathfrak{N}_0 = \mathfrak{H}^1 \cap \mathfrak{N}$, and $\mathfrak{H}^0 = \mathfrak{T}_0 \cap \mathfrak{H}^1$, so that $\mathfrak{N}_0 \subset \mathfrak{H}^0$. If $\mathfrak{N} = \mathfrak{N}_0$, then $\mathfrak{N} = 1$, for $\mathfrak{H}^1$ contains no non-trivial normal subgroups of $\mathfrak{H}^2$. Also notice that $\mathfrak{N}_0$ is the intersection of all the conjugates of $\mathfrak{H}^0$ by elements of $\mathfrak{H}^1$.

Now $\mathfrak{H}^0$ must have order two, since it consists of the elements of $\mathfrak{H}^1$ which interchange some other specified set of 2 letters (i.e., fixes one other system, but not both its letters). Thus, by hypothesis, $\mathfrak{H}^0$ is normal in $\mathfrak{H}^1$. But then $\mathfrak{N}_0 = \mathfrak{H}^0 \neq 1$, and so $\mathfrak{N} \neq \mathfrak{N}_0$. Hence $|\mathfrak{N}_0| = 2$, $|\mathfrak{N}| = 4$, and $\mathfrak{N} = \mathfrak{T}_0$.

$\mathfrak{N}_0$ is not normal in $\mathfrak{H}^2$, and has in fact, $n/2$ distinct conjugates (one for each of the systems) in $\mathfrak{H}^2$. But all the conjugates of $\mathfrak{N}_0$ are in $\mathfrak{N}$, and since $\mathfrak{N}$ has order 4, and thus contains at most three different subgroups of order two, we have $n/2 \leqq 3$, and so $n = 4$ or 6.

Now we return to our 3-designs. $\mathfrak{J}^3$ is a group with the properties of $\mathfrak{H}^2$ in Theorem 4.1, if $\mathfrak{J}^3$ is transitive on m_3, and so $k=3$ or 5. Thus $\mathfrak{J}^1-\mathfrak{J}^3$ must consist, respectively, of the square-roots of unity or of the fourth-roots of unity in the field $\mathcal{F}$. We finally show that both these situations automatically give rise to a group like $\mathfrak{H}^2$.

THEOREM 4.2. *If $\mathfrak{J}^1$ consists of the square-roots of unity, or of the fourth-roots of unity, then $\mathfrak{J}^3$ has order* 8 *or* 24, *respectively, and* $|\mathfrak{J}^3:\mathfrak{J}^2|=4$ *or* 6, *respectively. In all other cases where $\mathfrak{J}^1$ is not the multiplicative group of a subfield, we have* $|\mathfrak{J}^3:\mathfrak{J}^2|=2$.

Proof. We need only show the first sentence. Consider the mapping $A: x\rightarrow(x-1)/(x+1)$, where of course we are assuming that the field does not have characteristic two. Then on the elements $0, 1, -1, \infty$, A acts as the four-cycle $(0\ -1\ \infty\ 1)$ and so when $\mathfrak{J}^1$ has order two, A is in $\mathfrak{J}^3$ and $\mathfrak{J}^3$ is transitive. Furthermore, if $\mathcal{F}$ contains a fourth-root of unity, it is not difficult to see that A fixes the fourth-root of unity. So if $\mathfrak{J}^1$ has order 4, A is still in $\mathfrak{J}^3$ and again $\mathfrak{J}^3$ is transitive. This finishes the proof of the theorem.

So in case 2 we have three subcases:

Case 2a. If $k=3$, then π^3 is a 3-design for $(q+1,4,3)$, with $q(q^2-1)/8$ blocks; this case can only rise for q odd.

Case 2b. If $k=5$, then π^3 is a 3-design for $(q+1,6,5)$, with $q(q^2-1)/24$ blocks; this case can only arise for $q\equiv 1 \pmod 4$.

Case 2c. If $k\neq 3$ or 5, then π^3 is a 3-design for

$$(q+1,k+1,k(k+1)/2),\ \text{with}\ q(q^2-1)/2(k-1)\ \text{blocks}.$$

It is easy to compute that the 3-designs above are trivial only in the following cases: in case 2a, when $q=5$, and in case 2c, when $q=7$, $k=4$.

Without giving here a definition of either an affine or inversive plane, we remark that a finite affine plane is exactly the same as a 2-design for $(m^2,m,1)$, and a finite inversive plane is exactly the same as the extension of an affine plane, that is, is a 3-design for $(m^2+1,m+1,1)$. Hence in case 1 (when $q=p^2$) we have an infinite family of finite inversive planes; they are in fact the finite Miquelian ones, exactly. The 2-designs of case 1 are the finite affine spaces (Desarguesian in the plane case), but the 3-designs for $n>2$ are, in general, new to the author.

Now we can apply Theorem 3.1 in another way (in fact in several others), by choosing a subgroup of $\mathfrak{G}^2$ to define a block, and then using

Theorem 3.1 once, instead of twice. This situation has been partially studied in [8], and we report on the results here. Let m_2 be the set of points of a subgroup $\mathfrak{J}$ of $\mathcal{J}^*$, of order k. Then m_3 will consist of the points of m_2, plus ∞, and we report on the 3-designs obtained in this way:

Case 3a. $k=p-1$, where p is a prime-power and $q=p^n$. Then we have a 3-$(p^n+1, p, p-2)$, with $p^{n-1}(p^{2n}-1)/(p-1)$ blocks.

Case 3b. $k=3$, where q is odd. Then we have a 3-$(q+1, 4, 2)$, with $q(q^2-1)/12$ blocks.

Case 3c. Not 3a or 3b. Then we have a 3-$(q+1, k+1, k^2-1)$, with $q(q^2-1)/k$ blocks.

There are some other cases to examine, obviously, but we will not deal with them here. However there is another feature of these 3-designs which bears examining. Are any of them 4-designs in disguise, as it were? If this is to be so, then λ must satisfy a relation $\lambda=\lambda_3=\lambda_4(v-3)/(k-3)$, where k is the number of points on a block. There are infinitely many possibilities, it seems, and the author has checked a small finite number of these. Leaving aside certain trivial or uninteresting cases, there are two which are t-designs for $t>3$. In case 2c, with $q=16$, $k=6$, the design is in fact a 4-design for $(17,7,6)$, and in case 3c, with $q=11$, $k=5$, the design is even a 5-design for $(12,6,2)$. This 5-design, by the way, is only the third 5-design ever found, to the author's knowledge, and it bears a peculiar relationship to one of the other known 5-designs, which we discuss in the next section. (Of course, using the Mathieu group on 24 letters, it is possible to construct quite a few more non-trivial 5-designs with $v=24$.)

The details of the above facts will be found in [8], including a discussion of how the three subcases 3a, 3b, 3c arise.

5. Mathieu and Suzuki groups. We will consider the Mathieu groups first, but for our purposes it will suffice to know only a little about them, beyond the fact that they exist. A very satisfactory treatment of these groups is given by Witt in [12], and see also [4].

Let $\mathfrak{G}^2$ be the group of all transformations $x\to ax^{\alpha_a}+b$, where a, b, x are in $GF(9)$, $a\neq 0$, and where α_a is defined by $\alpha_a=1$ or 3 according as a is or is not a square in $GF(9)$. If $\mathfrak{G}^1$ is the subgroup of $\mathfrak{G}^2$ fixing 0, then $\mathfrak{G}^1$ is the quaternion group of order 8. By choosing a block to be the subgroup of order two, $\mathfrak{G}^1$ defines a 1-design for $(8,2,1)$, and since the element $x\to x+1$ is in $\mathfrak{J}^2$, it is immediate that $\mathfrak{G}^2$ gives us a 2-design for $(9,3,1)$; that is, the affine plane of order 3. It is well-known that $\mathfrak{G}^2$ has a transitive extension $\mathfrak{G}^3$ on 10 letters, and $\mathfrak{J}^3$ can be computed for this extension and

seen to be greater than $\mathfrak{J}^2$, whence $\mathfrak{G}^3$ defines a 3-design for $(10, 4, 1)$; that is, an inversive plane (in fact the only one with those parameters). Then, what is not at all obvious, $\mathfrak{G}^3$ has a transitive extension to a group $\mathfrak{G}^4$ on 11 letters, and $\mathfrak{G}^4$ has a transitive extension to a group $\mathfrak{G}^5$ on 12 letters; these last two groups are usually called the Mathieu groups $\mathfrak{M}_{11}$ and $\mathfrak{M}_{12}$. We can, using the specific permutations σ_i, for $i = 4$ and 5, as given by Witt in [12], compute $\mathfrak{J}^4$ and $\mathfrak{J}^5$. But we can use Theorem 3.5 to avoid this, even to avoid the computations connected with $\mathfrak{J}^3$. For $\mathfrak{G}^2$ is 2-flag-transitive, and since $\mathfrak{J}^1$ is characteristic in $\mathfrak{G}^1$, it is certainly $\mathfrak{G}^0$-special (here it happens that $\mathfrak{G}^0 = 1$). Thus we can conclude that any transitive extensions of $\mathfrak{G}^2$, no matter how many times carried out, must have $\lambda = 1$. The designs associated with $\mathfrak{G}^3$, $\mathfrak{G}^4$, $\mathfrak{G}^5$ then must have, respectively, 30, 66, and 132 blocks. If $\mathfrak{G}^5$ had a transitive extension $\mathfrak{G}^6$, then this would define a 6-design with $\lambda = 1$ and $(13)(132)/7$ blocks; this last number is no integer, and so we have:

THEOREM 5.1. *The groups* $\mathfrak{G}^3$, $\mathfrak{G}^4$, $\mathfrak{G}^5$ *above are* 3-, 4-, *and* 5-*flag-transitive collineation groups, respectively, of designs* 3-$(10, 4, 1)$, 4-$(11, 5, 1)$, *and* 5-$(12, 6, 1)$, *and* $\mathfrak{G}^5$ *cannot possess a transitive extension.*[3]

A related question, to which the author cannot supply an answer, is the following: what inversive planes have extensions? Using Corollary 1 to Theorem 3.1, it is immediate that the only interesting candidates, besides the one of order 3, are the ones of order 4, 8 and 13; for any example of order 2 is trivial, and it might be mentioned that an example of order 4 could not have a 4-transitive group, since there is no such non-trivial group on 18 letters.

Now we repeat this ad hoc analysis of the other Mathieu groups. Here we take $\mathfrak{G}^2$ to be the projective unimodular group of the plane over $GF(4)$, and so $\mathfrak{G}^2$ is presented to us in a natural way as a 2-flag-transitive collineation group on a 2-design for $(21, 5, 1)$; the design is even symmetric. In order to apply Theorem 3.5, we must know that $\mathfrak{J}^1$ is a $\mathfrak{G}^0$-special subgroup of $\mathfrak{G}^1$. Now $\mathfrak{G}^1$ is the subgroup of $\mathfrak{G}^2$ fixing a point, $\mathfrak{G}^0$ is the subgroup of $\mathfrak{G}^1$ fixing another point, and $\mathfrak{J}^1$ can be taken as the subgroup of $\mathfrak{G}^1$ fixing the line which joins these two points.

In dual form, $\mathfrak{G}^1$ is the subgroup fixing a line c, $\mathfrak{G}^0$ is the subgroup fixing two lines, c and d, while $\mathfrak{J}^1$ is the subgroup fixing the point P which is the intersection of c and d. If we think of $\mathfrak{G}^1$ as the group acting on the

[3] It is the 5-design here that was mentioned in the previous section. The 5-$(12, 6, 2)$ found there has the property that its blocks can be broken into two equal subsets such that the set of 12 points together with the blocks of either subset form a 5-$(12, 6, 1)$, and in fact form the 5-$(12, 6, 1)$ associated with the Mathieu group $\mathfrak{G}^5$.

affine plane obtained by deleting c, then it is completely straight-forward to see that if $\mathfrak{T}$ is the normal subgroup of $\mathfrak{G}^1$ made up of translations, than $\mathfrak{J}^1 = \mathfrak{T}\mathfrak{G}^0$. But in fact, $\mathfrak{T}$ is even characteristic in $\mathfrak{G}^1$, and therefore $\mathfrak{J}^1$ is $\mathfrak{G}^0$-special. Then, given that the extensions $\mathfrak{G}^3$, $\mathfrak{G}^4$, $\mathfrak{G}^5$ exist, we can say, in complete analogy with Theorem 5.1:

THEOREM 5.2. *The extensions $\mathfrak{G}^3$, $\mathfrak{G}^4$, $\mathfrak{G}^5$ of $\mathfrak{G}^2$ above, are 3-, 4-, and 5-flag-transitive collineation groups, respectively, of designs* 3-(22, 6, 1), 4-(23, 7, 1), *and* 5-(24, 8, 1), *and $\mathfrak{G}^5$ cannot possess a transitive extension.*

Proof. The first part of the theorem follows if we know that the groups exist. The designs will have, respectively, 77, (11)(23), and (3)(11)(23) blocks. If $\mathfrak{G}^5$ had a transitive extension, then it would define a 6-design with $\lambda = 1$ and (3)(11)(23)(25)/9 blocks; this is no integer, so the theorem is proved.

Witt found all the designs for the Mathieu groups in [12], but he used a somewhat different method to construct the designs for the groups on 22, 23, and 24 letters, specifically the method we have generalized in Theorem 3.4. Given that $\mathfrak{G}^5$ exists, we will look for an appropriate S-subgroup of G^0, fixing more than 5 letters. But G^0 contains a Sylow 1-subgroup of order 16, and this subgroup, which we can call B, is exactly the group of translations of the plane, with respect to a certain line. Thus B is normal in G^0 and fixes 5 points in the plane, and thus fixes 8 points when considered as a subgroup of G^5. So G^5 defines a 5-design for (24, 8, 1).

It is known that the Mathieu group $\mathfrak{M}_{11}$, which we normally think of as a 4-transitive group on 11 letters, also has a 3-transitive representation on 12 letters. If we say that $\mathfrak{G}^3$ is the group, of order 12.11.10.6, acting on 12 letters, then one can discover that $\mathfrak{G}^2$ is a simple group of order 660 and $\mathfrak{G}^1$ is a simple group of order 60. $\mathfrak{G}^0$ is a non-abelian group of order 6, and acts on 9 letters in the following way: it has two orbits, one of length 6 on which it acts regularly, and the other of length 3 on which it acts as a Frobenius group (that is, as the symmetric group on 3 letters). If $\mathfrak{T}$ is a Sylow 2-subgroup of $\mathfrak{G}^0$, then clearly $\mathfrak{T}$ acts on the 9 letters in four orbits of length 2 and one fixed letter. Thus from Theorem 3.4, $\mathfrak{G}^3$ defines a 3-design for (12, 4, 3), with 165 blocks.

There is in fact another 3-design associated with this representation of the Mathieu group, and we sketch an explanation. We construct first a 2-design π^2 for (11, 5, 2) by means of the difference set of multiplicative squares, mod 11; that is, the points of π^2 are the marks of $GF(11)$ and the blocks are the sets $\mathfrak{D} + x$, where $\mathfrak{D}$ is the set of five multiplicative non-zero squares in $GF(11)$. This well-known design has the mappings $x \rightarrow ax + b$

as collineations, where b is any element of $GF(11)$ and a is in $\mathfrak{D}$. This set of collineations is transitive, and the subgroup fixing 0 has two orbits; 2 and 5 are in different orbits. Now the following mappings are collineations of π^2:

$$(0)(1)(2\,9\,6)(3\,4\,7)(5\,8\,10),\ (0)(1)(2\,5)(3)(4\,7)(6\,8)(9\,10).$$

These mappings fix both 0 and 1, and by means of them, 2 can be sent to 5, so the collineation group of π^2 is doubly-transitive on π^2. They generate a subgroup of order 6 fixing 0 and 1, and it can be shown that together with the mappings $x \rightarrow ax + b$, where a is in $\mathfrak{D}$, these mappings generate the entire group $\mathfrak{G}^2$ of collineations of π^2. Thus the group has order $11.10.6 = 660$.

It remains to see that $\mathfrak{G}^2$ can be extended. The following permutation can play the role of σ_8 in conditions (1), (2), (3) of Section 3 (that is, Witt's theory):

$$(\infty\ 0)(1)(2\,8)(3)(4)(5\,6)(7)(9\,10).$$

Put another way, the mapping above generates a new set of blocks, using the new symbol ∞, where of course ∞ is adjoined to all the old blocks. But it is easy to see that where before we had 11 blocks, we now have 22, and $\mathfrak{G}^2$ sends the new 11 blocks into themselves. So $\mathfrak{G}^3$ exists and must have order $12.11.10.6$; it is 3-flag-transitive on a symmetric 3-design for $(12, 6, 2)$.

If we know that $\mathfrak{G}^3$ is simple, then from considerations of order we will know that it is abstractly isomorphic to $\mathfrak{M}_{11}$, since the orders are small. We omit the calculations by which this is shown, but basically it is only a matter of showing that $\mathfrak{G}^1$ is simple, and this is not difficult.

We finally mention the Suzuki groups. If $\mathfrak{G}^2$ is the Suzuki group $\mathfrak{S}(q)$ of order $q^2(q^2+1)(q-1)$, 2-transitive on q^2+1 letters, then $\mathfrak{G}^1$, of order $q^2(q-1)$, is Frobenius, with a Frobenius kernel of order q^2. This Frobenius kernel is a 2-group and its commutator subgroup has order q. We make $\mathfrak{G}^1$ define a 1-design by choosing for a block the letters corresponding to the elements of the commutator subgroup of the Frobenius kernel. Then the 1-design has parameters $(q^2, q, 1)$. From Theorem 3.1, $\mathfrak{G}^2$ defines a 2-design for $(q^2+1, q+1, \lambda)$. Since $\mathfrak{G}^1$ is flag-transitive, $\lambda = 1$ or $\lambda = q+1$. But $\lambda = 1$ means that the 2-design has $q(q^2+1)/(q+1)$ blocks, and this is no integer; so $\lambda = q+1$. But this 2-design is in fact a 3-design for $(q^2+1, q+1, 1)$, as a good deal of laborious calculation will reveal, and hence is an inversive plane. This inversive plane is not Miquelian, and these are in fact the only non-Miquelian finite inversive planes known at the present time. (Tits found these inversive planes from the Suzuki groups by different methods; see [11].)

Added in proof: In a paper ("Extension of designs and groups: projective, symplectic and certain affine groups," *Mathematische Zeitschrift*, vol. 89 (1965), pp. 199-205) the author has used methods developed from the material given here to prove some results about the non-existence of certain transitive extensions. In particular, that paper gives yet another proof that the Mathieu group of degree 24 has no transitive extension. Also, the definition of S-subgroup as given in this paper has been modified and sharpened in the *Mathematische Zeitschrift* paper.

WESTFIELD COLLEGE,
(UNIVERSITY OF LONDON)
LONDON N. W. 3.

REFERENCES.

[1] R. A. Beaumont and R. P. Peterson, "Set-transitive permutation groups," *Canadian Journal of Mathematics*, vol. 7 (1955), pp. 35-42.

[2] Peter Dembowski, "Kombinatorische Eigenschaften endlicher Inzidenzstrukturen," *Mathematische Zeitschrift*, vol. 75 (1960/61), pp. 256-270.

[3] ——— and D. R. Hughes, "On inversive planes with an orthogonality," *Journal of the London Mathematical Society*, vol. 40 (1965), pp. 171-182.

[4] M. Hall, Jr., *The Theory of Groups*, New York, 1958.

[5] D. G. Higman and J. W. McLaughlin, "Geometric ABA groups," *Illinois Journal of Mathematics*, vol. 5 (1961), pp. 382-397.

[6] D. R. Hughes, "Combinatorial analysis; t-designs and permutation groups," *Proceedings of the Symposium on Pure Mathematics*, vol. VI (1962), pp. 39-41.

[7] ———, "On k-homogeneous groups," *Archiv der Mathematik*, vol. 15 (1964), pp. 401-403.

[8] ———, "Some new t-designs for $t = 3$, 4, and 5," *Rendiconti di Matematica e delle sue Applicazioni*, to appear.

[9] H. J. Ryser, *Combinatorial Mathematics*, Carus Mathematical Monograph no. 14, 1963.

[10] M. Suzuki, "Contributions to the theory of finite groups," *Proceedings of the Symposium on Pure Mathematics*, vol. VI (1962), pp. 107-109.

[11] J. Tits, "Ovoïdes à translations," *Rendiconti di Matematica e delle sue Applicazioni* (5), vol. 21 (1962), pp. 37-59.

[12] E. Witt, "Die 5-fach transitiven Gruppen von Mathieu," *Abhandlungen aus dem Mathematischen Seminar Hamburg Universität*, vol. 12 (1937), pp. 256-264.

[13] ———, "Über Steinersche Systeme," *ibid.*, vol. 12 (1937), pp. 265-275.

[14] Hans Zassenhaus, "Über transitive Erweiterungen gewisser Gruppen aus Automorphismen endlicher mehrdimensionaler Geometrien," *Mathematische Annalen*, vol. 111 (1935), pp. 748-759.

SECANT BUNDLES ON SYMMETRIC PRODUCTS.

By Arthur Mattuck.*

A linear system $\mathfrak{L}$ on an algebraic curve X determines an invertible sheaf $\mathcal{L}$ on the curve, and the formulation of geometric properties of linear systems in terms of these sheaves is today one of the cohomological facts of life. Some of the finer properties of linear systems however can really only be expressed by considering $X(n)$, the n-fold symmetric product of the curve. In this paper accordingly we study a vector bundle $\mathcal{E}_n(\mathcal{L})$ of rank n on $X(n)$ which is derived from $\mathcal{L}$ by a symmetrization process, and we relate some of its properties to the linear system. The main geometric result is the determination of rational equivalence class on $X(n)$ of that cycle $\mathfrak{L}^{(n)}$ which represents the totality of positive divisors of degree n contained in the system $\mathfrak{L}$. This cycle is, or should have been, a classical object; it enters implicitly into various classical results. To illustrate these notions, if $\mathcal{L}$ is the contangent bundle to X, then $\mathcal{E}(\mathcal{L})$ is the cotangent bundle to $X(n)$, and $\mathfrak{L}^{(n)}$ is the locus of special divisors on $X(n)$. These loci were introduced in [2] and played an important role there; it was an endeavor to understand this role that was the starting point of the present paper.

In general, $\mathfrak{L}^{(n)}$ is one of the Chern classes of $\mathcal{E}_n(\mathcal{L})$, so the paper is in two parts: the study of $\mathcal{E}_n(\mathcal{L})$ in sufficient detail to calculate its Chern classes, and then the connection of these classes with $\mathfrak{L}^{(n)}$. On the one hand, the study of the structure of the bundle rests upon the happy notion of a "secant bundle," recently introduced by Schwarzenberge [7]; $\mathcal{E}_n(\mathcal{L})$ turns out to be a secant bundle in his sense. Already in a preliminary version of his paper was an explicit, though somewhat tentative, discussion of the connection between $\mathfrak{L}^{(n)}$ and secant bundles; the present paper in that sense continues this work. On the other hand, the connection of $\mathfrak{L}^{(n)}$ with Chern classes comes from two sources. First, the geometric interpretation of a Chern class as the locus over which a suitable set of global sections become dependent. This is known if there are enough sections to give an imbedding into the Grassmannian; if not, one can use the work of Porteus [5], and we give a brief account of what we need of it here, with some amplifications. Second, one needs to know the relation between sections of $\mathcal{L}$ and sections

Received July 7, 1964.

* Supported in part by a grant from the National Science Foundation.

of $\mathcal{E}_n(\mathcal{L})$; for characteristic p this is not entirely trivial, and in fact was one of the cornerstones of a recent proof of the Riemann-Roch theorem for curves [4]: we borrow it from that paper.

MacDonald [1] has studied $\mathfrak{L}^{(n)}$ in characteristic zero, by determining first the homology ring of $X(n)$ and proceeding from there to get $\mathfrak{L}^{(n)}$ without introducing the bundles $\mathcal{E}_n(\mathcal{L})$. The elegant formulations of the classical formulas all appear for the first time in his work. The approach in this paper is by contrast algebraic and "functorial," valid in arbitrary characteristic, and well-adapted for the study of families of curves; in addition the formulas are finer, being placed in the rational equivalence ring, rather than the homology ring.

The language of this paper is approximately that of Borel-Serre [8], with a few additions from Grothendieck's "Elements." In addition we assume Schwarzenberger's study of the symmetric product [6] and the first few pages of [4], though a brief summary of these is given which should give an adequate idea of their contents. A word about notation: $\mathfrak{A}(V)$ is the rational equivalence ring of a non-singular variety V, and if $f\colon X \to Y$ is a map of varieties, then f_* and f^* denote, according to the context, any of the forwards and backwards functorial maps associated with f: on the sheaves, the cycles, or the rational equivalence rings.

1. The symmetrization of an invertible sheaf. Our starting point is an irreducible non-singular curve X over an algebraically closed ground field k, its n-fold symmetric product $X(n)$, and its n-fold Cartesian product $X[n]$. We use $p_i\colon X[n] \to X$ for the i-th projection, and $\pi\colon X[n] \to X(n)$ for the natural Galois covering map, whose group G is the symmetric group on n letters. The points of $X(n)$ represent the positive divisors of degree n on X; if $\mathfrak{a}$ is such a divisor, we will also let $\mathfrak{a}$ denote the corresponding point of $X(n)$.

Let $\mathcal{L}$ be a fixed invertible (locally free of rank one) sheaf on X. Then $\mathcal{L}[n] = \oplus\, p_i^*\mathcal{L}$ is a locally free sheaf of rank n on $X[n]$, and from it we define the sheaf $\mathcal{E}(\mathcal{L})$ on $X(n)$ whose sections over any open U are the G-invariant sections of $\mathcal{L}[n]$ over $\pi^{-1}(U)$:

$$(1) \qquad \Gamma(U, \mathcal{E}(\mathcal{L})) = \Gamma(\pi^{-1}(U), \mathcal{L}[n])^G.$$

It is this sheaf $\mathcal{E}(\mathcal{L})$ that is the object of study in this paper. We will call it the *symmetrization* of $\mathcal{L}$.

For example, let Ω_Z denote as usual the contangent sheaf on a non-singular variety Z—i. e., the sheaf of germs of holomorphic 1-forms, or simple

differentials of the first kind. Then an important special case of a sheaf $\mathcal{E}(\mathcal{L})$ was given in [4], where it was proved that

$$\mathcal{E}(\Omega_X) = \Omega_{X(n)}. \tag{2}$$

Since it is well-known that $\Omega_X[n] = \Omega_{X[n]}$, the nontrivial statement embodied in (2) is that every G-invariant holomorphic 1-form on $\pi^{-1}(U)$ is the lifting of a *holomorphic* 1-form on U. This in turn is proved by using a differential analogue—valid in all characteristics—of the Newton identities of classical algebra.

We return now to the general case, fixing $\mathcal{L}$ once and for all and writing simply $\mathcal{E}$ for $\mathcal{E}(\mathcal{L})$. The cohomological study of $\mathcal{E}$ rests on identifying it with a certain "secant bundle" in the sense of [7], and this is Proposition 1 below. The relation of $\mathcal{E}$ to classical geometrical questions rests on Proposition 2, which we deduce trivially from (2) above.

PROPOSITION 1. *Consider the natural maps*

$$X \xleftarrow[q]{} X \times X(n-1) \xrightarrow[f]{} X(n), \qquad f(x, \mathfrak{a}) = x + \mathfrak{a}.$$

Then there is a natural isomorphism $\phi\colon f_*q^*\mathcal{L} \to \mathcal{E}(\mathcal{L})$.

Proof. We reduce it to an affine statement, and then translate into modules. Let V be an open affine subset of X, with coordinate ring $A = \Gamma(V, \mathcal{O}_X)$. Then $V[n]$ and $V(n)$ are affine open subsets of $X[n]$ and $X(n)$ respectively, and their coordinate rings we denote by $A[n]$ and $A(n)$. Thus $A[n]$ is the n-fold tensor product over k of A with itself, and $A(n)$ is the subring left fixed by the group G. Note that the sets of the form $V(n)$ cover $X(n)$, and the intersection of any two of them is again a set of the same type, since $V_1(n) \cap V_2(n) = W(n)$, where $W = V_1 \cap V_2$. To define ϕ, it is enough therefore to give for each $V(n)$ an isomorphism of the two $A(n)$-modules

$$\phi_{V(n)}\colon \Gamma(V(n), f_*q^*\mathcal{L}) \to \Gamma(V(n), \mathcal{E}) \tag{3}$$

whose restriction to $W(n)$ is $\phi_{W(n)}$.

We describe the right-hand module in (3). Consider the A-module $\Gamma(V, \mathcal{L})$, which we shall denote by L, and put

$$M_i = \Gamma(V[n], p_i^*\mathcal{L}) = A_1 \otimes \cdots \otimes A_{i-1} \otimes L \otimes A_{i+1} \otimes \cdots \otimes A_n$$
$$M = \Gamma(V[n], \mathcal{L}[n] = M_1 \oplus \cdots \oplus M_n.$$

Here the A_i are isomorphic copies of A, and it is clear that M_i and M are $A[n]$-modules. A permutation $\sigma \in G$ acting on $V[n]$ has an induced

action on $A[n]$ and on the $A[n]$-module M, giving a σ-linear isomorphism $\sigma\colon M_i \to M_{\sigma(i)}$. Thus $\sigma(am) = \sigma(a)\sigma(m)$, for $a \in A[n]$, $m \in M$. It follows that the right side of (3) is the $A(n)$-module

$$\Gamma(V(n), \mathcal{E}) = \Gamma(V[n], \mathcal{L}[n])^G = M^G.$$

On the other hand, since $f^{-1}(V(n)) = V \times V(n-1)$, we have for the left-hand side of (3),

$$\Gamma(V(n), f_* q^* \mathcal{L}) = \Gamma(V \times V(n-1), q^* \mathcal{L}) = L \otimes A(n-1).$$

This is a module over $A \otimes A(n-1)$, which is to be viewed as an $A(n)$-module by the natural ring injection $A(n) \to A \otimes A(n-1)$.

Now view G as the permutations of $(1, \cdots, n)$ and let G_1 be the subgroup leaving 1 fixed. Let $\sigma_1, \cdots, \sigma_n$ be coset representatives for G modulo G_1, indexed so that $\sigma_i(1) = i$. Then for $m_1 \in M_1 = L \otimes A[n-1]$, define a trace map

$$\Phi\colon L \otimes A[n-1] \to M, \qquad \Phi(m_1) = \sigma_1(m_1) + \cdots + \sigma_n(m_n).$$

We claim that this map induces on the module $L \otimes A(n-1)$ the desired isomorphism $\phi_{V(n)}$

$$\phi_{V(n)}\colon L \otimes A(n-1) \to M^G. \tag{3'}$$

In the first place $\phi_{V(n)}$ is an injection since π is a Galois covering and therefore separable, and its image is evidently contained in M^G. To see that it is onto, let $m = m_1 + \cdots + m_n \in M^G$, where $m_i \in M_i$. Since m is in particular G_1-invariant, so is m_1, so that actually $m_1 \in L \otimes A(n-1)$. But m is also invariant under the automorphisms $\sigma_1, \cdots, \sigma_n$, so we have $m_i = \sigma_i(m_1)$, and thus $m = \phi_{V(n)}(m_1)$.

The isomorphism $\phi_{V(n)}$ defined above on $V(n)$ evidently commutes with the localization induced by a localizing map $A \to A_{\mathfrak{p}}$, so the maps $\phi_{V(n)}$ patch together over the intersections to give the isomorphism claimed by the proposition.

COROLLARY. *There is a canonical injection*

$$\phi_L\colon \Gamma(X, \mathcal{L}) \to \Gamma(X(n), \mathcal{E}).$$

If X is complete, it is an isomorphism.

Proof. We have by the proposition an isomorphism

$$\Gamma(\phi)\colon H^0(X \times X(n-1), q^* \mathcal{L}) \to H^0(X(n), \mathcal{E}).$$

The Kunneth formula gives

$$H^0(X\times X(n-1), q^*\mathcal{L}) = H^0(X,\mathcal{L}) \otimes_k H^0(X(n-1), \mathcal{O})$$

and if X is complete, then $H^0(X(n-1),\mathcal{O}) = k$.

PROPOSITION 2. 1. *$\mathcal{E}$ is a locally free sheaf of rank n.*

2. *Let $\mathfrak{a}$ be a positive divisor of degree n on X, whose support is contained in an open set V. Let $s\in\Gamma(V,\mathcal{L})$, and put $s'=\phi_L(s)$. Then*

$$(s)\geqq\mathfrak{a} \Longleftrightarrow s'(\mathfrak{a})=0. \tag{4}$$

Remark. By (s) we mean $s^{-1}(0)$, the divisor of zeros of s, when $\mathcal{L}$ is viewed as a line bundle. By $s'(\mathfrak{a})=0$, we understand that s' is to be viewed as a section of the vector bundle associated with $\mathcal{E}$; in terms of the sheaf $\mathcal{E}$ therefore, it means that $s'_{\mathfrak{a}}\in m_{\mathfrak{a}}\,\mathcal{E}_{\mathfrak{a}}$, where $m_{\mathfrak{a}}$ is the maximal ideal in the local ring $\mathcal{O}_{\mathfrak{a}}$.

Proof. When $\mathcal{L}=\Omega_X$, we have the relation (2); the map ϕ_L is nothing but the usual symmetrization map sending a holomorphic 1-form on X into a holomorphic 1-form on $X(n)$. Evidently in this case therefore $\mathcal{E}(\mathcal{L})$ is locally free and (4) gives a connection between 1-forms on X and on $X(n)$ which was proved in [4]. The general case follows easily from this. We remark first that for an invertible sheaf $\mathcal{L}$ on X, and any divisor $\mathfrak{a}$ on X, one can always find an open $V\subset X$ containing $\mathrm{supp}(\mathfrak{a})$ and such that $\mathcal{L}\,|_V\cong\mathcal{O}\,|_V$. One can, for example, take $V=X-\mathrm{supp}(\mathfrak{b})$, where $\mathfrak{b}$ is any (not necessarily positive) divisor in the divisor class defining $\mathcal{L}$ whose support does not meet $\mathrm{supp}(\mathfrak{a})$. It follows that we can find a V containing $\mathrm{supp}(\mathfrak{a})$ and such that there is an $\mathcal{O}_V$-isomorphism $g_V\colon\mathcal{L}\,|_V\to\Omega_V$. This isomorphism extends naturally to the respective symmerizations over $V(n)$:

$$g_{V(n)}\colon \mathcal{E}(\mathcal{L})\,|_{V(n)}=\mathcal{E}(\mathcal{L}\,|_V)\to\mathcal{E}(\Omega_V)=\mathcal{E}(\Omega_X)\,|_{V(n)}=\Omega_{V(n)}$$

and it commutes with the map ϕ_L of the corollary (taking $X=V$). Since the assertions of the proposition are local in character, their truth for Ω implies their truth for $\mathcal{L}$.

2. The structure of $\mathcal{E}(\mathcal{L})$. We assume from now on that X is complete, of genus g. Our goal is to elucidate the structure of $\mathcal{E}(\mathcal{L})$, and this will be based on the interpretation of $X(n)$ given in [6], whose salient facts we now recall.

Let J be the Jacobian of X. Once a reference point $c_0\in X$ is fixed, there is a map $h_n\colon X(n)\to J$, where h (nc_0) is the identity of J; h_n is

uniquely determined (up to automorphism of J). The idea is now to give a sheaf on J whose associated (dual) projective bundle is $X(n)$; it is then this sheaf, rather than $X(n)$ itself, that will be central.

We have maps

$$J \underset{\alpha_x}{\overset{p}{\leftrightarrows}} J \times X \underset{\beta_y}{\overset{q}{\rightleftarrows}} X$$

where p, q are projections, $\alpha_x\colon J \to J \times x$ and $\beta_y\colon X \to y \times X$. On $J \times X$ there is an invertible sheaf $\mathcal{P}$ (the one associated with a Poincare divisor) which shows how J parametrizes the invertible sheaves on X of degree 0; formally, its basic property is that if $y = h(\mathfrak{a})$, then $\beta_y{}^*\mathcal{P}$ is the invertible sheaf $[\mathfrak{a} - nc_0]$ on X, i. e., the one associated with the divisor class to which $\mathfrak{a} - nc_0$ belongs. $\mathcal{P}$ is normalized by requiring also that $\alpha_{c_0}{}^*\mathcal{P} = \mathcal{O}_J$.

The fiber $h^{-1}(y)$ is the derived projective space of the vector space $H^0(X, [\mathfrak{a}])$, so by the Serre duality theorem, it is the derived dual projective space of the vector space $H^1(X, [\mathfrak{k} - \mathfrak{a}])$, where $[\mathfrak{k}] = \Omega_X$. What we want therefore is a sheaf on J whose 'fiber" over the point y is this last vector space. To this end, let $\mathcal{J}_m = [mc_0]$ on X, and put

$$\mathcal{F}_m = R^1 p_* (\mathcal{P} \otimes q^* \mathcal{J}_m). \tag{5}$$

Then by the exchange property (see below), the fiber $\mathcal{F}_m \otimes \kappa(y)$ of $\mathcal{F}_m$ over the point y is $H^1(X, \beta_y{}^*(\mathcal{P} \otimes q^* \mathcal{J}_m)$, which in turn equals

$$H^1(X, [\mathfrak{a} + (m - n)c_0])$$

by the basic property of $\mathcal{P}$. This is almost what we need: we have only to adjust the degrees and move the base by the usual map $\theta\colon J \to J$ defined by $\theta(y) = -y + c$, where c is the canonical point, $c = h_{2g-2}(\mathfrak{k})$. The final result is an isomorphism

$$X(n) \xrightarrow{\ \sim\ } \boldsymbol{P}(\theta^* \mathcal{F}_{2g-2-n})$$

where $\boldsymbol{P}$ is Grothendieck's "dual projective bundle functor" [9, II, 4.1]. This isomorphism commutes with the maps of both sides into J, and depends only on the choice of reference point c_0. Moreover, it makes the fundamental sheaf $\mathcal{O}(1)$ correspond to the invertible sheaf on $X(n)$ associated with the divisor $i(X(n-1))$, where i is the map

$$i\colon X(n-1) \to X(n), \qquad i(\mathfrak{a}) = \mathfrak{a} + c_0. \tag{6}$$

We will analyse $\mathcal{E}(\mathcal{L})$ by relating it to sheaves on J. There are two

steps. First we assume n is large, and then we get a result for smaller n by proving a theorem which is the classical adjunction formula in the special case $\mathcal{L}=\Omega_X$. In addition to standard cohomological tools, we use a theorem of Grothendieck, in the following not too general version:

"PRINCIPLE OF EXCHANGE." *Let $f: X \to Y$ be a proper morphism of two algebraic varieties, and suppose for all $y \in Y$, $\dim f^{-1}(y) \leqq p$. Let $g: Y' \to Y$ be an arbitrary base extension by another variety, set $X' = X \times_Y Y'$, $f': X' \to Y'$ and $g': X' \to X$ the projections. Then, provided either* (i) $p=0$, *or* (ii) $\mathfrak{F}$ *is coherent and flat over Y, we have for a quasi-coherent $\mathcal{O}_X$-Module $\mathfrak{F}$ a canonical isomorphism in the top dimension $g^*R^pf_*(\mathfrak{F}) \cong R^pf_*'(g'^*(\mathfrak{F}))$.*

For the first case, $p=0$, f is evidently a covering, and therefore affine, and the result is entirely elementary [9, II, 1.5.2]. The second case can be deduced from [9, III, 7.7.5] by deciphering the language and applying [9, III, 4.2.2]. We shall actually use this case only when f is a projection $Z \times Y \to Y$, where Z is projective; for this case the principle could be proved directly using affine coverings of Z and the Kunneth formula (note that the principle is local on Y and Y').

The basic commutative diagram for what follows is

$$\begin{tikzcd}
& X(n-1)\times X \arrow[dl, "f"'] \arrow[d, "j"] \arrow[dr, "q"] & \\
X(n) & X(n)\times X \arrow[l, "p'"] \arrow[r, "q'"'] & X
\end{tikzcd}$$

where the horizontal arrows are projections, f is as in Proposition 1, and j is the imbedding

$$(7)\qquad j: X(n-1)\times X \to X(n) \times X, \qquad j(\mathfrak{a},x)=(\mathfrak{a}+x,x).$$

Let W be the image of j—it is a non-singular subvariety of codimension 1, the graph of the standard correspondence between X and $X(n)$—and let $\mathfrak{J}(W)$ be the invertible $\mathcal{O}_{X(n)\times X}$-ideal defining W. We note that the map $p'|_W: W \to X(n)$ is a covering, hence is affine, and is essentially the same as f.

PROPOSITION 3. *If $n > \deg \mathcal{L}$, there is a canonical exact sequence of sheaves on $X(n)$*

$$0 \to \mathcal{O}_{X(n)}{}^{r_0} \to \mathcal{E}(\mathcal{L}) \to R^1p'_*(q'^*\mathcal{L}\otimes \mathfrak{J}(W)) \to \mathcal{O}_{X(n)}{}^{r_1} \to 0.$$

where $r_i = \dim H^i(X,\mathcal{L})$, $i=0,1$.

Proof. This comes from a sequence introduced in [7] by Schwarzenberger to analyse general secant bundles. The usual short sequence for the subvariety W of $X(n) \times X$ gives, after tensoring with $q'^*\mathcal{L}$,

$$(8) \qquad 0 \to q'^*\mathcal{L} \otimes \mathfrak{I}(W) \to q'^*\mathcal{L} \to q'^*\mathcal{L} \otimes j_*\mathcal{O}_{X(n-1)\times X} \to 0.$$

The exact sequence for the cohomological functor $R^i p'_*$, applied to (8), then gives the proposition. Examining the terms one by one, and starting from the right we have

(i) $R^i p'_*(q'^*\mathcal{L} \otimes j_*\mathcal{O}) = R^i f_*(q^*\mathcal{L})$. If $i = 1$, this gives 0 since f is a covering. If $i = 0$, we get $\mathcal{E}(\mathcal{L})$ by Proposition 1.

(ii) $R^i p'_*(q'^*\mathcal{L}) = \mathcal{O}_{X(n)}{}^{r_i}$ since the module of sections of this sheaf over an affine $U \subset X(n)$ is by definition

$$H^i(U \times X, p'^*\mathcal{O}_{X(n)} \otimes q'^*\mathcal{L}) = H^i(X, \mathcal{L}) \otimes \mathcal{O}_U.$$

(iii) If $y = \mathfrak{a}$ is any point of $X(n)$, and $g\colon X \to y \times X$ the corresponding closed imbedding, then by definition of W, we have $g^*\mathfrak{I}(W) = [-\mathfrak{a}]$. Therefore, $g^*(q'^*\mathcal{L} \otimes \mathfrak{I}(W)) = \mathcal{L} \otimes [-\mathfrak{a}]$, which has no sections if $n > \deg \mathcal{L}$, so that $H^0(p'^{-1}(y), q'^*\mathcal{L} \otimes \mathfrak{I}(W) \otimes \kappa(y)) = 0$ for all y; since $q'^*\mathcal{L} \otimes \mathfrak{I}(W)$ is flat over $X(n)$, we conclude by [9, III, 4.6.1] that $p'_*(q'^*\mathcal{L} \otimes \mathfrak{I}(W)) = 0$.

PROPOSITION 4. *Let* $\mathcal{L} = [\mathfrak{a}]$, $\deg \mathfrak{a} = l$, *and*[1] $h(\mathfrak{a}) = a \in J$. *Define* $\theta_a\colon J \to J$ *by* $\theta_a(y) = -y + a$. *Then for all* $n \geqq 0$, *there is an isomorphism*

$$R^1 p'_*(q^*\mathcal{L} \otimes \mathfrak{I}(W)) \to h^*\theta_a^*(\mathfrak{F}_{l-n}) \otimes \mathcal{O}(-1).$$

Proof. The diagram which is relevant is now

$$\begin{tikzcd} J \times X \arrow[d, "p"'] & X(n) \times X \arrow[l, "h'"'] \arrow[d, "p'"] \\ J & X(n) \arrow[l, "h"] \end{tikzcd} \qquad h' = h \times 1_X$$

There is a relation connecting divisors on $X(n) \times X$:

$$W \sim h'^*(P) + n(X(n) \times c_0) + i(X(n-1) \times X$$

where P is a Poincare divisor; it follows from the basic property of P and

[1] For any divisor $\mathfrak{a}$ on X, we may define $h(\mathfrak{a}) = h_m(\mathfrak{a}') - h_n(\mathfrak{a}'')$, where $\mathfrak{a} = \mathfrak{a}' - \mathfrak{a}''$, $\mathfrak{a}' \geqq 0$, $\mathfrak{a}'' \geqq 0$, $\deg \mathfrak{a}' = m$, $\deg \mathfrak{a}'' = n$.

the "see-saw principle." Changing the signs throughout and translating into sheaves gives [6, prop. 9]

$$\mathcal{A}(W) = h'^*(\mathcal{P}^{-1} \otimes q^*\mathcal{J}_{-n}) \otimes p'^*\mathcal{O}(-1).$$

Since $\mathcal{O}(-1)$ is locally free, we can apply [8, § 5(d)], and thus get

$$\begin{aligned} R^1p'_*(q^*\mathcal{L} \otimes \mathcal{A}(W)) &= R^1p'_*h'^*(\mathcal{P}^{-1} \otimes q^*(\mathcal{J}_{-n} \otimes \mathcal{L}) \otimes \mathcal{O}(-1) \\ &= h^*R^1p_*(\mathcal{P}^{-1} \otimes q^*(\mathcal{J}_{-n} \otimes \mathcal{L})) \otimes \mathcal{O}(-1) \end{aligned}$$

by the principle of exchange. By an easy calculation, for any m,

$$(\theta_a \times 1_X)^*(\mathcal{P} \otimes q^*\mathcal{J}_m) = \mathcal{P}^{-1} \otimes q^*(\mathcal{J}_{m-l} \otimes \mathcal{L}).$$

Applying a trivial principle of exchange gives therefore, using (5),

$$R^1p_*(\mathcal{P}^{-1} \otimes q^*(\mathcal{J}_{-n} \otimes \mathcal{L}) = \theta_a^*R^1p_*(\mathcal{P} \otimes q^*\mathcal{J}_{l-n}) = \theta_a^*\mathcal{F}_{l-n}.$$

Example. Propositions 3 and 4 give the structure of $\mathcal{E}(\mathcal{L})$ if $n > \deg \mathcal{L}$. To relate this to something more familiar, take $\mathcal{L} = \Omega_X$, so that $\mathcal{E}(\mathcal{L}) = \Omega_{X(n)}$. Combining these propositions gives us for $n > 2g-2$ an exact sequence on $X(n)$.

$$0 \to \mathcal{O}_{X(n)}{}^g \xrightarrow[\alpha]{} \Omega_{X(n)} \to h^*\theta^*\mathcal{F}_{2g-2-n} \otimes \mathcal{O}(-1) \xrightarrow[\beta]{} \mathcal{O}_{X(n)} \to 0.$$

Now Ω_J is a trivial sheaf of rank g on J, so that $h^*\Omega_J \cong \mathcal{O}_{X(n)}{}^g$. Letting $\mathcal{K}$ be the cokernel of the map α therefore, the above sequence decomposes into

$$\text{(i)} \qquad 0 \to h^*\Omega_J \xrightarrow[\alpha]{} \Omega_{X(n)} \to \mathcal{K} \to 0$$

$$\text{(ii)} \qquad 0 \to \mathcal{K} \to h^*\theta^*\mathcal{F}_{2g-2-n} \otimes \mathcal{O}(-1) \xrightarrow[\beta]{} \mathcal{O}_{X(n)} \to 0.$$

The sequence (i) is the standard exact sequence for the projective fiber bundle $h\colon X(n) \to J$ relating the cotangent sheaves to the base space and the bundle space with the "cotangent bundle along the fibers," $\mathcal{K}$. One has to check from the relevant part of Proposition 3 that α is indeed the natural injection, under the identification of $h^*\Omega_J$ with $\mathcal{O}_{X(n)}{}^g$. The sequence (ii) gives the explicit formula $\mathcal{K} = \mathcal{H} \otimes \mathcal{O}(-1)$, where $\mathcal{H}$ is defined by

$$0 \to \mathcal{H} \to h^*\theta^*\mathcal{F}_{2g-2-n} \to \mathcal{O}(1) \to 0.$$

This however is the standard formula for the cotangent sheaf along the fibers to the projective bundle $\boldsymbol{P}(\mathcal{B})$, where $\mathcal{B}$ is a locally free sheaf over some base variety. in our case, $\theta^*\mathcal{F}_{2g-2-n}$ is locally free, since $n > 2g-2$ [6, prop. 3].

We now complete Propositions 3 and 4 by studying $\mathcal{E}(\mathcal{L})$ for lower values of n. Since n will vary, we use subscripts $\mathcal{E}_n = \mathcal{E}(\mathcal{L})$, $\mathcal{O}_n = \mathcal{O}_{X(n)}$, etc. when the n must be specified.

PROPOSITION 5. *If i_n is the imbedding* (6), *then there is an exact sequence on* $X(n-1)$

$$0 \to \mathcal{O}_{n-1}(-1) \to i_n{}^*\mathcal{E}_n \to \mathcal{E}_{n-1} \to 0.$$

Proof. For simplicity, if T is a divisor on a variety V, whose corresponding $\mathcal{O}_V$-Ideal is $\mathfrak{J}(T)$, then as is often done we will denote the quotient $\mathcal{O}_V/\mathfrak{J}(T)$ by $\mathcal{O}_T$ (instead of $\alpha_*\mathcal{O}_T$, where $\alpha\colon T \to V$ is the inclusion).

The geometric picture is

$$\begin{tikzcd} Y \arrow[r, "j'"] \arrow[d, "i''_n"] & X(n-1)\times X \arrow[r, "p_{n-1}"] \arrow[d, "i'_n"] & X(n-1) \arrow[d, "i_n"] \\ X(n-1)\times X \arrow[r, "j_n"'] & X(n)\times X \arrow[r, "p_n"'] & X(n) \end{tikzcd} \qquad \begin{tikzcd} X(n-1)\times X \arrow[r, "q"] \arrow[d, "i'_n"] & X \arrow[d] \\ X(n)\times X \arrow[r, "q_n"'] & X. \end{tikzcd}$$

Here j_n is the map (7), i_n is the map (6), $i'_n = i_n \times 1_X$, Y is the fiber product of j_n and i'_n, with projections i''_n and j'.

Put $W_n =$ image of j (as in Proposition 3), and set $Z = X(n-1) \times c_0$. Then viewing these as cycles on $X(n)\times X$ and $X(n-1)\times X$, respectively, we have the simple cycle formulas

(i) $\qquad j'_*Y = i'_n{}^*(W_n) = W_{n-1} + Z$

(ii) $\qquad W_{n-1}\cdot Z = i_{n-1}(X(n-2)) \times c_0$.

In $X(n-1)\times X$ there is an exact sequence

$$(9) \qquad 0 \to \mathcal{B} \to \mathcal{O}_Y \to \mathcal{O}_{W_{n-1}} \to 0, \qquad \mathcal{B} \equiv \mathfrak{J}(W_{n-1})/\mathfrak{J}(Y).$$

Since we are dealing with divisors on non-singular varieties, we have by (i)

$$\mathfrak{J}(Y) = \mathfrak{J}(W_{n-1})\cdot\mathfrak{J}(Z) \qquad \mathfrak{J}(W_{n-1}\cdot Z) = \mathfrak{J}(W_{n-1}) + \mathfrak{J}(Z).$$

Thus by one of the basic isomorphism theorems, there is an exact sequence (at first on $X(n)\times X$)

$$(10) \qquad 0 \to \mathcal{B} \to \mathcal{O}_Z \to \mathcal{O}_{W_{n-1}\cdot Z} \to 0.$$

Evidently this sequence is still exact when viewed on Z. Moreover, the map p_{n-1} induces an isomorphism of Z with $X(n-1)$. Let $p = p_{n-1}$, $q = q_{n-1}$, and apply the functor $p_*(q^*\mathcal{L}\otimes\ )$ to (10); this gives

$$0 \to p_*(q^*\mathcal{L}\otimes\mathcal{B}) \to \mathcal{O}_{X(n-1)} \to \mathcal{O}_{i_{n-1}(X_{n-2})} \to 0,$$

so that $p_*(q^*\mathcal{L}\otimes\mathcal{B}) = \mathfrak{J}(i_{n-1}(X(n-2))) = \mathcal{O}_{n-1}(-1)$ by (6).

Applying the same functor to (9) gives exactly the exact sequence of the proposition. The first term we have just checked. The fourth term would be $R^1p_*(q^*\mathcal{L}\otimes\mathcal{B})$, which is 0 since p restricted to $Z=\operatorname{supp}(q^*\mathcal{L}\otimes\mathcal{B})$ is a covering. Referring back to Proposition 1, we have $f_{n-1}=p_{n-1}j_{n-1}$ so that proposition gives for the third term of our sequence

$$p_*(q^*\mathcal{L}\otimes\mathcal{O}_{W_{n-1}})=f_*q^*\mathcal{L}=\mathcal{E}_{n-1}(\mathcal{L}).$$

Finally, for the second term, using (i) we have

$$\begin{aligned} p_*(q^*\mathcal{L}\otimes\mathcal{O}_Y) &= p_*(q^*\mathcal{L}\otimes i'^*\mathcal{O}_{W_n}) = p_*i'^*(q_n{}^*\mathcal{L}\otimes\mathcal{O}_{W_n}) \\ &= p_*i'^*j_*(j^*q^*\mathcal{L}). \end{aligned}$$

But in general, for quasi-coherent sheaves we have $p_*i'^*j_*=p_*j'_*i''=i^*p_{n*}j_*$, using the principle of exchange (case i) first for the affine morphism j, then for the affine morphism p_nj. The above therefore becomes

$$\begin{aligned} &= i^*p_{n*}j_*(j^*q^*\mathcal{L}) \\ &= i^*p_{n*}(q^*\mathcal{L}\otimes\mathcal{O}_{W_n}) \\ &= i^*\mathcal{E}_n \qquad \text{by Proposition 1.} \end{aligned}$$

3. The Chern classes of $\mathcal{E}_n(\mathcal{L})$. We have enough information now to determine the Chern classes of $\mathcal{E}_n(\mathcal{L})$, but first we must describe the elements of the rational equivalence rings $\mathfrak{A}(X(n))$ and $\mathfrak{A}(J)$ which enter into the formulas. In general, these elements carry subscripts denoting their degree (i.e., codimension); c_0 is the reference point on X.

Looking first at $\mathfrak{A}(X(n))$, there is a map

$$j^r\colon X(n-r)\times X\to X(n), \qquad\qquad j^r(\mathfrak{a},z)=\mathfrak{a}+rz.$$

If $\mathfrak{b}$ is a divisor (or divisor class) on X, we then set

$$\begin{aligned} \xi_r(\mathfrak{b}) &= j^r{}_*(1\times\mathfrak{b}), \qquad\qquad \xi=\xi_1(c_0) \\ \delta_{r-1} &= j^r{}_*(1\times 1). \end{aligned}$$

The element ξ is the divisor class in $\mathfrak{A}(X(n))$ corresponding to the sheaf $\mathcal{O}(1)$. Moreover, it is easily seen that $\xi^r=\xi_r(c_0)$, as in [2, § 4].

Let the maps $h\colon X(n)\to J$ and $\theta_a\colon J\to J$ be as in Proposition 4; note that h_* lowers degree by $n-g$ and that θ_a depends on $\mathcal{L}$. Set, in $\mathfrak{A}(J)$,

$$w_i=h_*(\xi^{n-g+i}), \qquad\qquad w_i'=\theta_a{}^*(w_i),$$

Finally, we recall from [2] or [6] the determination of the total Chern class of $\mathcal{F}_s$, for $s<g$:

$$c(\mathcal{F}_s)=1+w_1+w_2+\cdots+w_g. \tag{11}$$

This may be proved either by direct calculation (as in [2]), or better (as in [6]) in the following two steps. First, by repeated application of an exact cohomology sequence, one gets an exact sequence on J

$$0 \to \theta^* \mathfrak{F}_{2g-2-n}{}^* \to \mathfrak{M} \to \mathfrak{F}_s \to 0, \qquad s < g, n > 2g-2$$

where $\mathfrak{M}$ is a succession of extensions by $\mathcal{O}_J$ and $*$ denotes the dual bundle. This gives $c(\mathfrak{F}_s)c(\theta^*\mathfrak{F}_{2g-2-n}{}^*) = c(\mathfrak{M}) = 1$. On the other hand, since $\mathcal{O}(1) = [\xi]$, a simple formal property of the Chern classes of any locally trivial derived projective bundle yields, using the above definition of w_i, $c(\theta^*\mathfrak{F}_{2g-2-n}{}^*)(1 + w_1 + \cdots + w_g) = 1$, from which therefore (11) follows.

THEOREM 1. *Let* $\deg \mathcal{L} = l$, *and set* $d = n + g - 1 - l$. *Then for all* $n > 0$, *the total Chern class of* $\mathcal{E}_n(\mathcal{L})$ *is given in* $\mathfrak{A}(X(n))$ *by*

$$c(\mathcal{E}_n(\mathcal{L})) = (1-\xi)^d + (h^*w_1')(1-\xi)^{d-1} + \cdots + (h^*w_g')(1-\xi)^{d-g}.$$

Remark. In using he formula, note that $(1-x)^{-i} = [(1-x)^{-1}]^i$, and $(1-x)^{-1} = 1 + x + x^2 + \cdots + x^n$, for any $x \in \mathfrak{A}(X(n))$.

Proof. For $n > l$, we have by Propositions 3 and 4, and the above,

$$c(\mathcal{E}_n) = c[h^*\theta_a{}^*(\mathfrak{F}_{l-n}) \otimes \mathcal{O}(-1)], \qquad c(\mathcal{O}(-1)) = 1 - \xi.$$

This gives the result immediately, in view of (11) and the usual formula for the Chern class of a tensor product; note that rank $\mathfrak{F}_{l-n} = d$, [6, prop. 3].

For smaller values of n, we use descending induction. Introduce n as an index as before: $h = h_n$, $i = i_n$ (the map (6)), and write $\xi^{(n)}$ for ξ. Then $h_n i_n = h_{n-1}$ and $i_n{}^*\xi^{(n)} = \xi^{(n-1)}$ (see, e.g. [2, § 4]). Then Proposition 5 gives

$$c(\mathcal{E}_{n-1}) = c(i_n{}^*\mathcal{E}_n)(1 - \xi^{(n-1)})^{-1} = i_n{}^*c(\mathcal{E}_n)(1 - \xi^{(n-1)})^{-1}.$$

But

$$i_n{}^*[(h_n{}^*w_1')(7 - \xi^{(n)})^m](1 - \xi^{(n-1)})^{-1} = (h_{n-1}{}^*w_1')(1 - \xi^{(n-1)})^{m-1}$$

so that if Theorem 1 is true for n, it is true for $n-1$ as well.

COROLLARY 1. *The above formula gives the Chern classes of* $X(n)$, *as* $c_i(X(n)) = (-1)^i c_i(\mathcal{E}_n(\Omega_X))$.

Proof. This follows from (2); the signs are changed since $\Omega_{X(n)}$ is the dual of the tangent bundle.

COROLLARY 2. *Viewed as elements of the numerical equivalence ring* $\mathfrak{N}(X(n))$,

$$c(\mathcal{E}_n(\mathcal{L})) = (1-\xi)^d e^{u(1-\xi)^{-1}}, \qquad u = h^*w_1.$$

Proof. We have in $\mathfrak{N}(J)$ the formula $w_i' = w_i = w_1^i/i!$, from which the formula follows, [3].

In a preliminary version of [7], Schwarzenberger gave a very different calculation which yields directly the exponential Chern classes of $\mathcal{E}_n(\mathcal{L})$ as elements of $\mathcal{A}(X(n)) \otimes Q$. It does not use the Jacobian. The resulting formulas are not formally deducible from the above by using known relations in $\mathcal{A}(X(n))$ or $\mathcal{A}(J)$; thus the equivalence of these two results gives new relations in both of these rings. These relations in $\mathcal{A}(J)$ connect various Pontrjagin multiples and intersections of the w_i; they are rather complicated and we do no more here than advertise their existence.

PROPOSITION 6 [Schwarzenberger]. *If* $\mathfrak{k} = c_1(\Omega_X)$, $\alpha = c_1(\mathcal{L})$, *then*

$$(-1)^i i!\, ch_i(\mathcal{E}_n(\mathcal{L})) = \frac{\delta_i}{i+1} - \xi_i(\alpha + (\frac{i-1}{2})\mathfrak{k}).$$

Proof. We have as in Proposition 3,

$$\begin{tikzcd}
 & W_n \arrow[dl, "q"'] \arrow[d, "j"] \arrow[dr, "f"] & \\
X & X(n) \times X \arrow[l, "q'"] \arrow[r, "p"'] & X(n)
\end{tikzcd}$$

where $W_n \cong X(n-1) \times X$ and j is thought of this as an inclusion map. Let ch and td denote respectively the exponential Chern class and the Todd class; then by the generalized Riemann-Roch theorem [8]—note that $f_! = f_*$ since f is affine)

$$ch(\mathcal{E}_n)\, td X(n) = f_*(q^* ch\mathcal{L} \cdot td\, W_n).$$

If $\mathfrak{N}$ denotes the normal bundle to W in $X(n) \times X$, then since W is non-singular, we have on $\mathcal{A}(W)$

$$td\, W \cdot td\, \mathfrak{N} = j^* td(X(n) \times X) = q^* td\, X \cdot f^*\, td\, X(n),$$

so that

$$ch(\mathcal{E}_n) = f_*(q^*(ch\, \mathcal{L} \cdot td\, X)\, (td\, \mathfrak{N})^{-1}).$$

Now $c(\mathcal{L}) = 1 + \alpha$, $c(X) = 1 - \mathfrak{k}$, $c(\mathfrak{N}) = 1 + \eta$ (say), $td\, X = 1 - \mathfrak{k}/2$, and $(td\, \mathfrak{N})^{-1} = \eta^{-1}(1 - \exp(-\eta)) = \Sigma(-\eta)^i/(i+1)!$, whence

$$(-1)^i i!\, ch_i(\mathcal{E}_n) = f_*(\frac{\eta^i}{i+1} - \eta^{i-1} q^*(\alpha - \mathfrak{k}/2))$$

Now η is the class of $j^{-1}(W)$ in $\mathcal{A}_1(W)$. A straightforward calculation

(proceeding mostly on $X[n]\times X$, along the lines of [2, § 4]) yields the desired result.

Evidently the numerical equivalence class of $\xi_i(z)$ is independent of z; denoting it by ξ_i we get therefore in the numerical equivalence ring

COROLLARY. $(-1)^i i! ch_i(\mathcal{E}_*(\mathcal{L}))=\frac{\delta_i}{i+1}-[(i-1)(g-1)+l]\xi_i.$

This is MacDonald's result, proved in characteristic zero for the case $\mathcal{L}=\Omega_X$ (see [1, (15.16)]).

4. A geometric interpretation of Chern classes. In what follows we abandon all previous notation and consider vector bundles over a fixed non-singular base variety X. We will not distinguish between a locally free sheaf of finite rank and its corresponding vector bundle. Also, o always denotes the zero section, viewed as a subvariety or a cycle, of whatever bundle one is looking at. If f is a map between varieties, f^* and f_* denote the corresponding maps on cycles or on the rational equivalence rings, or on the sheaves, according to the context.

Suppose then E is a rank p vector bundle, and let $s\colon X\to E$ be a (regular) section whose image in E intersects o properly. Then $s^{-1}(o)$ is of pure codimension p on X, and it is well-known that in $\mathfrak{A}(X)$ we have for the p-th Chern class

$$c_p(E)=s^*(o). \tag{12}$$

Porteous [5] has given an algebraic version of Thom's formulas connecting the singularities of a map with Chern classes; they imply among other things an analogue of (12) for the other Chern classes which we give here *ab ovo* for the convenience of the reader.

Definition. Let S be an r-dimensional subspace of $H^0(X,E)$. For each $x\in X$, S determines a subspace $S_x=\{s(x)\mid s\in S\}$ of the fiber E_x. The *dependency locus* $\mathfrak{D}(S)$ is defined to be $\{x\mid \dim S_x<r\}$.

We wish to define $\mathfrak{D}(S)$ not merely as a set, but also as a cycle, under some reasonable hypotheses. The simplest would be that it have the right dimension; actually we must assume here somewhat more.

Let P^{r-1} be the projective space (of lines through the origin) associated with S; if $s\in S$, $s\neq 0$, we will denote the corresponding point of P by $\bar{s}$. P carries a natural line bundle $\mathcal{O}(-1)$, which is a subbundle of the trivial bundle $P\times S$ over P. Let $\pi\colon P\times X\to X$ and $\rho\colon P\times X\to P$ be the projections. We now work over the space $P\times X$ and define a section σ of the

rank p bundle $\mathrm{Hom}(\rho^*\mathcal{O}(-1),\pi^*E)$ by $\sigma(\bar{s},x)[s]=s(x)$, and a closed subset of $P\times X$ by

$$\mathfrak{D}'(S)\equiv\sigma^{-1}(o)=\{(\bar{s},x)\mid s(x)=0\}.$$

Evidently $\pi(\mathfrak{D}'(S))=\mathfrak{D}(S)$, which justifies in part the

Definition. If $\mathfrak{D}'(S)$ is pure p-codimensional, the cycle on X defined by $D(S)=\pi_*\sigma^*(o)$ is called the *dependency cycle* of S.

PROPOSITION 7.

1. $\operatorname{cod}\mathfrak{D}(S)\leqq p-r+1$ *always*

2. *If* $D(S)$ *is defined, then* $\operatorname{supp}(D(S))=\mathfrak{D}(S)$, *and*

$$\operatorname{cod}D(S)=p-r+1.$$

Proof. 1. Working over a small open set U in X over which E becomes trivial, choose bases for the fiber and for S; then in terms of the bases, the moving subspace $S_x\subset E_x$ is given by an $r\times p$-matrix of functions regular on U, and $\mathfrak{D}(S)\cap U$ is the set where this matrix has rank $<r$. This is expressed locally on X by the vanishing of a suitable set of $p-r+1$ r-rowed minors of this matrix, so $\operatorname{cod}\mathfrak{D}(S)\leqq p-r+1$.

2. We have $\operatorname{supp}D(S)=\operatorname{supp}(\pi_*\sigma_0(o))\subset\pi(\operatorname{supp}\sigma^*(o))=\pi(\mathfrak{D}'(S))=\mathfrak{D}(S)$. To prove the reverse inclusion, let Z be a component of $\mathfrak{D}(S)$; since $\pi(\mathfrak{D}'(S))=\mathfrak{D}(S)$, Z is the image under π of some component Z' of the cycle $\sigma^*(o)$. By assumption, Z' has codimension p on $P\times X$; since π cannot raise dimension, Z must have codimension $\geqq p-r+1$. Therefore by part 1, $\operatorname{cod}Z=p-r+1$, $\dim Z'=\dim Z$, $\pi_*(Z')$ is some positive multiple of Z, and so $Z\subset\operatorname{supp}(\pi_*\sigma(o))$.

THEOREM 2. *Suppose that* $H^0(X,E)$ *has an* r*-dimensional subspace* S *whose dependency cycle* $D(S)$ *is defined. Then in* $\mathcal{A}(X)$, $c_{p-r+1}(E)=D(S)$.

In other words, the Chern class in question is interpreted geometrically as the locus where r sections become dependent—provided there *are* r sections, they are not *too* dependent, and you count multiplicities correctly.

Proof. From (12), we have in $\mathcal{A}(P\times X)$,

$$s^*(o)=c_p(\mathrm{Hom}(\rho^*\mathcal{O}(-1),\pi^*(E))).$$

Since $\mathcal{O}(-1)$ and $\mathcal{O}(1)$ are dual bundles, applying π_* to the above gives in $\mathcal{A}(X)$.

$$D(S)=\pi_*c_p(\rho^*\mathcal{O}(1)\otimes\pi^*E).$$

Now $c(\mathcal{O}(1))=1+h$, where h is the class in $\mathcal{A}(P)$ of a hyperplane. By the formula for the Chern class of a tensor product, the above

$$\begin{aligned}&=\pi_*[(\rho^*h)^p+c_1(\pi^*E)(\rho^*h)^{p-1}+\cdots+c_p(\pi^*E)]\\&=\pi_*\rho^*(h^p)+c_1(E)\pi_*\rho^*(h^{p-1})+\cdots+c_p(E)\pi_*\rho^*(h^0).\end{aligned}$$

But this last is just $c_{p-r+1}(E)$, since evidently $\pi_*\rho^*(h^i)=0$ unless $i=r-1$, in which case it equals 1.

5. Application to linear systems on curves. In the classical theory of curves, a positive divisor $\mathfrak{a}$ is called "special" if it is contained in the canonical class, that is, if there exists a positive divisor $\mathfrak{k}$ of the canonical class such that $\mathfrak{k}=\mathfrak{a}+\mathfrak{b}$ for some positive divisor $\mathfrak{b}$. More generally, given any linear system, one may consider the divisors of some fixed degree n contained in it, and view them as forming a closed subset of $X(n)$. In order to make sense of the classical formulas, this subset must be viewed as a cycle on $X(n)$; our subject is to determine the rational eqivalence class of this cycle.

Translating the foregoing into sheaves, to give a linear system $\mathfrak{L}$ of degree l and dimension r means to give an invertible sheaf $\mathcal{L}$ of degree l and a subspace $L\subset H^0(X,\mathcal{L})$ of dimension $r+1$, the two being connected by $\mathfrak{L}=\{(s)\mid s\in L\}$, where $(s)=s^*(o)$. This subspace determines a subspace $L^{(n)}=\phi_L(L)$ of $H^0(X(n),\mathcal{E}_n(\mathcal{L}))$, by the corollary to Proposition 1, also of dimension $r+1$.

PROPOSITION 8. *If $r\leqq n\leqq l$, then $D(L^{(n)})$ is defined, and its support consists of the points of $X(n)$ representing the divisors of degree n contained in $\mathfrak{L}$.*

Proof. If P is the projective r-space associated with L (or with $L^{(n)}$), we must look at

$$\mathcal{D}'(L^{(n)})=\{(\bar{s},\mathfrak{a})\mid s\in L,\phi_L(s)[\mathfrak{a}]=0\}=\{(\bar{s},\mathfrak{a})\mid(s)\geqq\mathfrak{a}\}$$

using Proposition 2. This shows however that $\mathcal{D}'$ is a finite covering of P, since over each point $\bar{s}$ of P lie the finite number of positive divisors $\mathfrak{a}$ of degree n contained in the divisor (s) on X. Therefore $\mathcal{D}'$ is pure r-dimensional, which shows by definition that $D(L^{(n)})$ is defined; since $\operatorname{supp}(D(L^{(n)}))=\pi(D')$, the rest follows immediately.

In view of Proposition 8 and the "naturalness" of its definition, we will say that the cycle $D(L^{(n)})$, which we shall abbreviate to $\mathfrak{L}^{(n)}$, represents the divisors of degree n contained in the system $\mathfrak{L}$. We have then, combining Theorems 1 and 2,

THEOREM 3. *For $r \leqq n \leqq l$, the rational equivalence of the cycle $\mathfrak{L}^{(n)}$ is given in $\mathcal{A}(X(n))$ as the term of degree $n-r$ in, equivalently,*

$$(1-\xi)^d + (h^*w_1')(1-\xi)^{d-1} + \cdots + (h^*w_g')(1-\xi)^{d-g}, \text{ or}$$
$$(1+\xi)^{l-r-g}(1+h^*w_1' + \cdots + h^*w_g').$$

The second expression given above is (up to numerical equivalence) MacDonald's form for the result. The equivalence of the two is formal. His definition of the cycle $\mathfrak{L}^{(n)}$ is formally different from the one used here, but may be shown to be equivalent. We omit the proof.

Two interesting special cases of the above formula are the following.

1. If the linear system $\mathfrak{L}$ is complete and nonspecial (i.e., contains no special divisors), then by the Riemann-Roch theorem $n-r=d+1$, and we get simply $\mathfrak{L}^{(n)}=h^*w_{n-r}'$. This equation is valid even for the cycles; it shows therefore how $\mathfrak{L}^{(n)}$ looks on $X(n)$, and shows also that it consists of a single irreducible component, with multiplicity one. More generally, if $r=l-g$, this equation is valid in $\mathcal{A}(X(n))$, though not as an equality of cycles.

2. If $\mathfrak{L}$ is the canonical system, then $d=n-g+1$ and

$$\mathfrak{L}^{(n)} = h^*w_d' - (h^*w_{d-1}')\xi + \cdots + (-1)^d\xi^d$$

so that

$$h^*w_d' = \mathfrak{L}^{(n)} + i_*\mathfrak{L}^{(n-1)} \qquad (i \text{ is the map } (6)).$$

This last is the basic formula (5) of [2]. It identifies $\mathfrak{L}^{(n)}$ with the cycle there called $S^{(n-g)}$, consisting of a single irreducible component counted with multplicity one and whose point set is the totality of points on $X(n)$ representing special divisors on X of degree n. Furthermore,

$$(-1)^n\mathfrak{L}^{(n)} = c_{n-g+1}(X(n)),$$

which "explains" why these particular subvarieties are important.

As an application of the theorem, we sketch a proof of the deJonquieres formula valid in arbitrary characteristic. This formula is proved for characteristic 0 in [1, (17.2)]; as a substitute for the detailed knowledge of the homology ring of $X(n)$ required there, we reduce the calculation to some intersection formulas in the numerical equivalence ring of J first given in the abstract case by Matsusaka and Weil.

As in [1], let $\varpi = (p_1^{n_1}, \cdots, p_k^{n_k})$ be a partition of n; so $n_i > 0$,

$\sum p_i n_i = n$, and the p_i are distinct positive integers. The totality of positive divisors $\mathfrak{a}$ on X of degree n which have the form

$$\mathfrak{a} = p_1\mathfrak{a}_1 + \cdots + p_k\mathfrak{a}_k, \qquad \deg \mathfrak{a}_i = n_i$$

consitute a subvariety $\Delta(\varpi)$ of dimension $n_1 + \cdots + n_k$ on $X(n)$.

Given a linear system $\mathfrak{L}$, of degree l and dimension r, the classical deJonquieres formula gives the number of divisors of degree n contained in this system which have the above form, provided this number is finite. Like many enumerative problems, there is some vagueness in the classical statements, it being tacitly assumed either that $\mathfrak{L}$ is fairly generic, or else that one counts multiplicities correctly. What is evidently called for is the intersection number $[\mathfrak{L}^{(n)} \cdot \Delta(\varpi)]$, assuming these two cycles have complementary dimension (i.e., that $n_1 + \cdots + n_k = n - r$) and that they intersect properly (which may not be the case in characterisic p, due to the failure of Bertini's theorem, for certain X and certain systems $\mathfrak{L}$).

The calculation proceeds in several steps. First, let $\xi^i = \xi_i(c_0)$ be as in § 3; then a standard sort of calculation on $X[n]$ yields the formula in $\mathcal{A}(X(n)) \otimes Q$ (cf. [2, § 4])

$$(13) \qquad \xi^i \cdot \Delta(\varpi) = \sum_{(J)} \binom{i}{j_1 \cdots j_k} p_1^{j_1} p_2^{j_2} \cdots p_k^{j_k} \Delta(p_1^{n_1 - j_1}, \cdots, p_k^{n_k - j_k})$$

where the sum is taken over all $(j_1, \cdots, j_k)$ for which $i = \sum j_\beta$, $0 \leq j_\beta \leq n_\beta$.

Next, let $C = h_1(X)$ be the generating curve of the Jacobian J, and let $p\delta$ denote the usual endomorphism of J (the p-th multiple of the identity), and $*$ be the Pontrjagin product in $\mathcal{A}(J)$. Then it is easily seen that in $\mathcal{A}(J)$,

$$h_*(\Delta(\varpi)) = (p_1\delta) W_{g-n_1} * (p_2\delta) W_{g-n_2} * \cdots * (p_k\delta) W_{g-n_k}$$

where $W_{g-m} = h_m(X(m))$. Now in the numerical equivalence ring $\mathcal{N}(J)$, we have [3] the relations $(p\delta) C \equiv p^2 C$, $C^{(*i)} = i! W_{g-i}$, from which we get $(p\delta) W_{g-i} \equiv p^{2i} W_{g-i}$. Applying this to the above gives

$$(14) \qquad h_*(\Delta(\varpi)) = p_1^{2n_1} \cdots p_k^{2n_k} \binom{m}{n_1, \cdots, n_k} W_{g-m}, \qquad m = \sum n_\beta.$$

Combining (13) and (14) gives

$$(15) \qquad h_*(\xi^i \cdot \Delta(\varpi)) \equiv [i, \varpi] W_{g-m+i}$$

where $[i, \varpi]$ is the coefficient of $t_1^{n_1} \cdots t_k^{n_k}$ in $(\sum p_\beta t_\beta)^i (\sum p_\beta^2 t_\beta)^{m-i}$. Finally, to caluclate $[\mathfrak{L}^{(n)} \cdot \Delta(\varpi)]$ and get the formula, we may calculate the

degree of $h_*(\mathfrak{L}^{(n)} \cdot \Delta(\varpi))$ instead. This is easily done by using (15) and Theorem 3, together with the formulas $[w_i \cdot w_{g-i}] = \binom{g}{i}$ and $w_i' \equiv w_i$ in $\mathfrak{N}(J)$ (see [3]). The result is finally, $(m = n - r)$,

$[\mathfrak{L}^{(n)} \cdot \Delta(\varpi)] =$ coefficient of $t_1^{n_1} \cdots t_k^{n_k}$ in $(1 + \sum p_\beta t_\beta)^{l-r-g}(1 + \sum p_\beta{}^2 t_\alpha)^g$ which, with all the *caveats* understood, gives the number of divisors of type ϖ contained in the linear system $\mathfrak{L}$ of degree l and dimension r.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] I. G. MacDonald, "Symmetric products of an algebraic curve," *Topology*, vol. I (1962), pp. 319-343.

[2] A. Mattuck, "Symmetric products and Jacobians," *American Journal of Mathematics*, vol. 83 (1961), pp. 189-206.

[3] ———, "On symmetric products of curves," *Proceedings of the American Mathematical Society*, vol. 13 (1962), pp. 82-87.

[4] A. Mattuck and A. Mayer, "The Riemann-Roch theorem for algebraic curves," *Annali della Scuola Normale Superiore di Pisa*, serie III (Mathematics) vol. XVII (1963), pp. 223-237.

[5] I. Porteus, 'Simple singularities of maps," to appear.

[6] R. L. E. Schwarzenberger, "Jacobians and symmetric products," *Illinois Journal of Mathematics*, vol. 7 (1963), pp. 257-268.

[7] ———, "The secant bundle of a projective variety," *Proceedings of the London Mathematical Society*, vol. 14 (1964), pp. 369-384.

[8] J. P. Serre and A. Borel, "Le theoreme de Riemann-Roch," *Bulletin de le Société Mathematique de France*, vol. 86 (1958), pp. 97-136.

[9] A. Grothendieck, "Elements de geometrie algebrique," *Publications Mathematiques Institut des Hautes Etudes Scientifiques*, vol. 8 (1961), 11, (1962), 17 (1963).

ON THE REDUCTION OF INDUCED INDECOMPOSABLE REPRESENTATIONS.

By PATRICIA A. TUCKER.

1. Introduction. Let G be a finite group with a normal subgroup H. In [5] and [6] an explicit method was given for splitting up the induced module $L^G = KG \otimes_{KH} L$ into components when L was assumed to be an irreducible left KH-module and K was algebraically closed. In this paper, the same method is used to handle some of the cases that occur when L is indecomposable and K is algebraically closed.

The method is limited in that one must use a K-factor set of a certain subgroup S of G/H. In general, the factor set is in $\mathrm{Hom}_{KH}(L, L)$ which is the semi-direct sum of its radical and $K \cdot 1_L$. Assuming that the factor set is in $K \cdot 1_L$, one can apply the construction given in [5] and [6] and prove a general intertwining theorem. In § 4, Theorem 3 gives some cases in which the construction can be applied and an example is given in which the factor set is not in $K \cdot 1_L$.

Furthermore, when the factor set is in K, the components can be written as modules induced from the tensor product of two modules over twisted group algebras for the inertia group of L (as in Theorem 3 of Clifford [1]). This can be applied to $G = G_1 \times G_2$ to show that an indecomposable left KG-module which is (G, G_2)-projective is an outer tensor product $L_1 \# L_2$ where L_i is an indecomposable left KG_i-module, $i = 1, 2$.

The basic definitions and notations used in this paper will be found in Curtis and Reiner [3] and in [5] and [6]. All modules are unital and finite dimensional vector spaces over K.

2. Construction and intertwining theorem. As in [6], let G be a finite group with a normal subgroup H and let $G/H = B = \{b_1 = 1, b_2, \cdots, b_n\}$. Then, $G = H \cup \bar{b}_2 H \cup \cdots \cup \bar{b}_n H$. Multiplication of elements of G is given by

$$\bar{b}h \cdot \bar{b}'h' = \overline{bb'}(b, b')h^{b'}h', \text{ where } b, b' \in B, h, h' \in H$$

and $h^b = (\bar{b})^{-1}h\bar{b}$ and $(b, b') \in H$. The elements (b, b') are called the factor set of the extension and satisfy

$$(bb', b'')(b, b')^{b''} = (b, b'b'')(b', b'').$$

Received July 20, 1964.

Let L be an indecomposable left KH-module over an algebraically closed field K. L affords a representation T of H. For each $g \in G$, $g \otimes L$ is a left KH-module since $h(g \otimes l) = g(g^{-1}hg) \otimes l = g \otimes h^g l$ and is indecomposable.

$S = \{b \in B \mid \bar{b} \otimes L$ is KH-isomorphic to $L\}$ is a subgroup of B. For each $b \in S$, select a non-singular linear transformation D_b of L such that $D_b h^b = hD_b$ and $D_1 = 1_L$. D_b determines a KH-isomorphism of $\bar{b} \otimes L$ onto L by $\bar{b} \otimes l \to D_b l$.

Let $a, b \in S$. Then,

$$D_a D_b (h^a)^b D_b^{-1} D_a^{-1} = h = D_{ab} h^{ab} D_{ab}^{-1}.$$

But $h^{ab} = (a,b)(h^a)^b (a,b)^{-1}$. Thus,

$$D_{ab}(a,b)(h^a)^b(a,b)^{-1} D_{ab}^{-1} = h.$$

Then, define $\rho(a,b) = D_a D_b (a,b)^{-1} D_{ab}^{-1}$ and $\sigma^b = D_b \sigma D_b^{-1}$. The following are easily verified.

LEMMA 1. $\rho(a,b) \in \mathrm{Hom}_{KH}(L,L)$

LEMMA 2. For $a, b, c \in S$,

$$\rho(b,c)^a \rho(a,bc) = \rho(a,b)\rho(ab,c).$$

If L is irreducible, then Schur's Lemma gives that $\rho(a,b) \in K \cdot 1_L$ and can thus be considered to be in K. Then, $b \to D_b$ is a projective representation of S. In general, $\mathrm{Hom}_{KS}(L^S, L^S)$ is the (left) crossed product of $\mathrm{Hom}_{KH}(L,L)$ and S with factor set $\rho(b,b')$ and correspondence given by $b \to b^*$ where $b^* : \sigma \to \sigma^b$. (See [7].)

Since K is algebraically closed and L is indecomposable, $\mathrm{Hom}_{KH}(L,L)$ is completely primary and can be written as

$$\mathrm{Hom}_{KH}(L,L) = K \cdot 1_L \oplus \mathrm{N}(iL) \tag{1}$$

where $N(iL)$ is the radical of $\mathrm{Hom}_{KH}(L,L)$ and $\oplus$ denotes semi-direct sum, i.e., direct sum as vector spaces over K. **Assume that** $\rho(a,b) \in K \cdot 1_L$, for every $a, b \in S$. Then, $\rho(a,b) = \beta(a,b) \cdot 1_L$, for some

$$\beta(a,b) \in K^* = \{k \in K \mid k \neq 0\}.$$

From Lemma 2, it follows that β is a factor set of S.

Then, as in [6], one uses $(KS)_{\beta^{-1}}$ to obtain the indecomposable components of L^G and Theorem 1 holds—summarized briefly below:

Let $B = S \cup a_2 S \cup \cdots \cup a_r S$. Define for $s = \sum_{b \in S} \xi_s(b) b \in (KS)_{\beta^{-1}}$

$$c(a_\gamma, s, l) = \sum_{b \in S} \xi_s(b) \overline{a_\gamma b} \otimes (a_\gamma, b) D_b^{-1} l.$$

Let $C(I)$ be the K-subspace of L^G generated by

$$\{c(a_\gamma, s, l) \mid 1 \leqq \gamma \leqq r, s \in I, l \in L\}.$$

Then, $(C(I):K) = [B:S](I:K)(L:K)$ and

THEOREM 1. *If $I_1, \cdots, I_t$ are left ideals of $(KS)_{\beta^{-1}}$ such that $(KS)_{\beta^{-1}} = I_1 \oplus \cdots \oplus I_t$, then $L^G = C(I_1) \oplus \cdots \oplus C(I_t)$.*

THEOREM 2. *Let I_1 and I_2 be left ideals of $(KS)_{\beta^{-1}}$. Then, there exists a K-homomorphism of $\mathrm{Hom}_{KG}(C(I_1), C(I_2))$ onto $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$. If $I_1 = I_2 = I$, then there exists an algebra isomorphism between*

$$\frac{\mathrm{Hom}_{KG}(C(I), C(I))}{N(iC(I))} \quad \text{and} \quad \frac{\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I, I)}{N(iI)}$$

where the denominators are the appropriate radicals.

Remark. A much stronger theorem holds in the irreducible case. See Theorem 2 of [6].

Proof. Suppose $\{s_{j1}\}$ is a basis for I_1 over K and $\{s_{i2}\}$ is a basis for I_2 over K. Then, for $t = 1, 2$,

$$\{c(a_\gamma, s_{it}, l_\mu) \mid 1 \leqq \gamma \leqq r, 1 \leqq i \leqq (I_t:K), 1 \leqq \mu \leqq (L:K)\}$$

is a basis for $C(I_t)$ over K. These will be ordered lexicographically by (γ, i, μ).

Let $A \in \mathrm{Hom}_{KG}(C(I_1), C(I_2))$. Let $A_{\gamma i\delta j}$ be the block of A determined by the rows (γ, i, μ) and the columns (δ, j, ν), μ, ν varying. Then,

$$A(c(a_\delta, s_{j1}, l)) = \sum_{\gamma, i} c(a_\gamma, s_{i2}, A_{\gamma i\delta j} l).$$

Since $Ah = hA$ and

$$\begin{aligned} Ah(c(a_\delta, s_{j1}, l)) &= \sum c(a_\gamma, s_{i2}, A_{\gamma i\delta j} h^{a_\delta} l) \\ hA(c(a_\delta, s_{j1}, l)) &= \sum c(a_\gamma, s_{i2}, h^{a_\gamma} A_{\gamma i\delta j} l), \\ A_{\gamma i\delta j} h^{a_\delta} &= h^{a_\gamma} A_{\gamma i\delta j}, \text{ for every } \gamma,\ \delta,\ i,\ j,\ h. \end{aligned}$$

Thus, $A_{\gamma i\delta j} \in \mathrm{Hom}_{KH}(\bar{a}_\delta \otimes L, \bar{a}_\gamma \otimes L)$. For $\gamma \neq \delta$, this set contains no isomorphisms. For $\gamma = \delta$, this set is identified with $\mathrm{Hom}_{KH}(L, L)$.

Furthermore, $\bar{a}_\delta A = A\bar{a}_\delta$, for every δ, and

$$\begin{aligned} A\bar{a}_\delta(c(1, s_{j1}, l)) &= A(c(a_\delta, s_{j1}, l)) = \sum c(a_\gamma, s_{i2}, A_{\gamma i\delta j} l) \\ \bar{a}_\delta A(c(1, s_{j1}, l)) &= \bar{a}_\delta(\sum c(a_\gamma, s_{i2}, A_{\gamma i1j} l)) \\ &= \sum c(a_\lambda, b \cdot s_{i2}, D_b(a_\lambda, b)^{-1}(a_\delta, a_\gamma) A_{\gamma i1j} l) \end{aligned}$$

where $a_\delta a_\gamma = a_\lambda b$. Consider the a_δ terms in each one (take $\gamma = \delta$ in the first equation; $\gamma = 1$, $\lambda = \delta$, $b = 1$ in the second equation). Thus,

$$c(a_\delta, s_{i2}, A_{\delta i\delta j}l) = c(a_\delta, s_{i2}, A_{1i1j}l)$$

so that $A_{\delta i\delta j} = A_{1i1j}$, for every i and j. Since

$$A_{1i1j} \in \mathrm{Hom}_{KH}(L, L) = K \cdot 1_L \oplus N(iL),$$

$$A_{1i1j} = \alpha_{ij} 1_L + N_{ij}, \text{ where } \alpha_{ij} \in K,\ N_{ij} \in N(iL). \tag{2}$$

Next consider $\bar{b}A = A\bar{b}$ for $b \in S$. For $t = 1, 2$, let $b \cdot s_{jt} = \sum_r f^t{}_{rj}(b) s_{rt}$. Then,

$$\begin{aligned} \bar{b}A(c(1, s_{j1}, l)) &= \bar{b}(\sum c(a_\gamma, s_{i2}, A_{\gamma i1j}l)) \\ A\bar{b}(c(1, s_{j1}, l)) &= A(c(1, b \cdot s_{j1}, D_b l)) \\ &= A(\sum f^1{}_{rj}(b) c(1, s_{r1}, D_b l)) \\ &= \sum f^1{}_{rj}(b) c(a_\gamma, s_{q2}, A_{\gamma q1r} D_b l). \end{aligned}$$

Equate terms with $\gamma = 1$.

$$\begin{aligned} \bar{b}(\sum c(1, s_{i2}, A_{1i1j}l)) &= \sum f^2{}_{qi}(b) c(1, s_{q2}, D_b A_{1i1j} l) \\ &= \sum f^1{}_{rj}(b) c(1, s_{q2}, A_{1q1r} D_b l). \end{aligned}$$

Thus,

$$\sum_i f^2{}_{qi}(b) D_b A_{1i1j} = \sum_r f^1{}_{rj}(b) A_{1q1r} D_b.$$

Then,

$$\sum_i f^2{}_{qi}(b) D_b A_{1i1j} D_b{}^{-1} = \sum_i f^1{}_{ij}(b) A_{1q1i}.$$

Using the decomposition (2) of the A_{1i1j}'s and the uniqueness of the decomposition (1), this gives

$$\sum_i f^2{}_{qi}(b) \alpha_{ij} 1_L = \sum_i f^1{}_{ij}(b) \alpha_{qi} 1_L \tag{3}$$

and

$$\sum_i f^2{}_{qi}(b) D_b N_{ij} D_b{}^{-1} = \sum_i f^1{}_{ij}(b) N_{qi}.$$

Now consider an element P of $\mathrm{Hom}_{(KS)_{g^{-1}}}(I_1, I_2)$. Suppose $P = [p_{ij}]$ so that $Ps_{j1} = \sum_i p_{ij} s_{i2}$. Then,

$$\begin{aligned} P(b \cdot s_{j1}) &= P(\sum_i f^1{}_{ij}(b) s_{i1}) = \sum_{i,q} f^1{}_{ij}(b) p_{qi} s_{q2} \\ b \cdot Ps_{j1} &= b \cdot (\sum_i p_{ij} s_{i2}) = \sum_{i,q} p_{ij} f^2{}_{qi}(b) s_{q2}. \end{aligned}$$

Thus,

$$\sum_i f^2{}_{qi}(b) p_{ij} = \sum_i f^1{}_{ij}(b) p_{qi}. \tag{4}$$

Therefore, (3) holds if and only if $P = [\alpha_{ij}]$ is an element of $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$. Thus, $\theta: A \to [\alpha_{ij}]$ is a mapping of $\mathrm{Hom}_{KG}(C(I_1), C(I_2))$ into $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$. From the uniqueness of the decomposition of $\mathrm{Hom}_{KH}(L, L)$, this is obviously a K-homomorphism.

Given $P \in \mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$, one can define $A \in \mathrm{Hom}_K(C(I_1), C(I_2))$ by $A_{\delta i \delta j} = p_{ij} 1_L$ and $A_{\gamma i \delta j} = 0$ for $\gamma \neq \delta$. Then, A is a KH-homomorphism since each $A_{\gamma i \delta j} \in \mathrm{Hom}_{KH}(\bar{a}_\delta \otimes L, \bar{a}_\gamma \otimes L)$. $\overline{a_\lambda b} A = A \overline{a_\lambda b}$ holds since (4) holds for P. Thus, A is a KG-homomorphism. Therefore, θ is a K-homomorphism onto.

Suppose that $I = I_1 = I_2$ and $s_i = s_{i1} = s_{i2}$. Then, θ will be an algebra homomorphism. Let $C = AB$. Then,

$$C_{1i1j} = \sum_{\lambda, k} A_{1i\lambda k} B_{\lambda k 1 j} = \Big(\sum_k \alpha_{ik} \beta_{kj} \Big) 1_L \qquad \mathrm{mod}\, N(iL)$$

so that $[\gamma_{ij}] = [\alpha_{ij}][\beta_{ij}]$. This follows since for $\lambda \neq 1$, $A_{1i\lambda k} B_{\lambda k 1 j}$ does not have an inverse and is thus in $N(iL)$.

In this case the kernel R of θ is a 2-sided ideal in $\mathrm{Hom}_{KG}(C(I), C(I))$, $= E$. $R = \{A \in E \mid A_{1i1j} \in N(iL)$, for every $i, j\}$. R includes A's such that each $A_{1i1j} = 0$ and some $A_{\gamma i \delta j} \neq 0$, for $\gamma \neq \delta$, but as noted previously the $A_{\gamma i \delta j}$'s for $\gamma \neq \delta$ are not isomorphisms. It is to be proven that $R \subseteq N(iC(I))$ $=$ radical of E.

Let $J_{\gamma i}$ be the K-subspace of L^G generated by $\{c(a_\gamma, s_i, l) \mid l \in L\}$. Since L is indecomposable, $J_{\gamma i}$ is an indecomposable left KH-module. Then,

$$C(I)_H = \sum_{\gamma, i} \oplus J_{\gamma i} \text{ (direct sum as } KH\text{-modules)}.$$

If $A \in E$, then $A \in \mathrm{Hom}_{KH}(C(I)_H, C(I)_H) = E_H$ and $A_{\gamma i \delta j}$ is a KH-homomorphism of $J_{\delta j}$ into $J_{\gamma i}$. By Theorem 8, p. 60, of [4], the radical of E_H is the set of $B \in E_H$ such that for every γ, δ, i, j, $B_{\gamma i \delta j}$ is not an isomorphism. Then, $R \subseteq$ radical of E_H. This means that R is a nilpotent 2-sided ideal of E. Therefore, $R \subseteq N(iC(I))$.

Since $E/R \cong \mathrm{Hom}_{(KS)_{\beta^{-1}}}(I, I)$, $N(iC(I))/R \cong N(iI)$ and thus

$$\frac{\mathrm{Hom}_{KG}(C(I), C(I))}{N(iC(I))} \cong \frac{\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I, I)}{N(iI)}.$$

COROLLARY 1. *I is indecomposable if and only if $C(I)$ is indecomposable.*

Proof. $\mathrm{Hom}_{KG}(C(I), C(I))$ is completely primary if and only if $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I, I)$ is completely primary.

COROLLARY 2. *L^G is indecomposable if and only if $(KS)_{\beta^{-1}}$ is indecom-*

posable, that is, if and only if $S=\{1\}$ or S is a p-group where p is the characteristic of K. (Last part is proved in [2] and [7].)

COROLLARY 3. *I_1 and I_2 are $(KS)_{\beta^{-1}}$-isomorphic if and only if $C(I_1)$ and $C(I_2)$ are KG-isomorphic.*

Proof. From the construction presented, $I_1 \cong I_2$ implies $C(I_1) \cong C(I_2)$. Conversely, if $C(I_1) \cong C(I_2)$, then there exists an invertible element in $\mathrm{Hom}_{KG}(C(I_1), C(I_2))$ and thus in $\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I_1, I_2)$. Hence, $I_1 \cong I_2$.

3. Further results. If $S \neq B$, then $C(I)$ is an induced module (see [5]). Let $\bar{S}H = \{\bar{b}h \mid b \in S, h \in H\}$. This is a subgroup of G. Define I' to be the K-subspace of L^G generated by $\{c(1,s,l) \mid s \in I, l \in L\}$. Then, I' is a left $K\bar{S}H$-module and $(I')^G$ is KG-isomorphic to $C(I)$.

Furthermore, I' is $K\bar{S}H$-isomorphic to the (inner) tensor product of two modules which afford projective representations of $\bar{S}H$. These are obtained as follows:

Define a factor set β^* of $\bar{S}H$ by $\beta^*(\bar{b}h, \bar{b}'h') = \beta(b, b')$. Then, let M be the left $(K\bar{S}H)_{\beta^*}$-module which is L with module operation given by $(\bar{b}h)l = D_b hl$. $(\beta^{-1})^* = (\beta^*)^{-1}$ is the factor set of $\bar{S}H$ given by

$$(\beta^{-1})^*(\bar{b}h, \bar{b}'h') = \beta(b, b')^{-1}.$$

Let $I(\bar{S}H)$ be the left $(K\bar{S}H)_{(\beta^{-1})^*}$-module which is I with module operation $\#$ given by $\bar{b}h \# s = b \cdot s$ where $\cdot$ denotes multiplication in $(KS)_{\beta^{-1}}$. Then, $I' \cong I(\bar{S}H) \otimes M$ so that $C(I) \cong (I(\bar{S}H) \otimes M)^G$.

These results extend Theorem 3 of Clifford [1] (or see [3], Theorem 51.7).

Remark. If I is indecomposable, then I' is indecomposable. This follows from the proof of Theorem 2 since $\mathrm{Hom}_{K\bar{S}H}(I', I')$ is determined by the A_{1l_1j}'s and thus

$$\frac{\mathrm{Hom}_{K\bar{S}H}(I', I')}{N(iI')} \cong \frac{\mathrm{Hom}_{(KS)_{\beta^{-1}}}(I, I)}{N(iI)}.$$

4. Application.

THEOREM 3. *Each of the following conditions imply that the construction presented can be performed.*

1) *L is irreducible.*

2) *G is a split extension of H by B and there exists a homomorphism*

of S into H such that $b \to h_b$ and $h_b h^b h_b^{-1} l = hl$, for every $h \in H$, $b \in S$, $l \in L$. This occurs if $h^b l = hl$, for every h, b, l.

3) *For every $b \in S$, there exists an $h_b \in H$ such that $h_b h^b h_b^{-1} l = hl$, for every h, b, l, and $h_b h_{b'} = h_{bb'}(b, b')$, for every $b, b' \in S$.*

4) *H is a complete group, i.e., all automorphisms of H are inner and its center is $\{1\}$.*

Proof of 2). Take each $D_b = h_b$. Since $b \to h_b$ is a homomorphism, $D_{bb'} = h_{bb'} = h_b h_{b'} = D_b D_{b'}$. Since the extension is split, each $(b, b') = 1$. Thus $\rho(b, b') = 1_L$.

Proof of 3). Take each $D_b = h_b$. Then $\rho(b, b') = 1_L$.

Proof of 4). $h \to h^b$ will be an inner automorphism of H determined by say $h_b \in H$. Then let $D_b = h_b$.

$$h_{bb'}{}^{-1} h h_{bb'} = \overline{\bar{b}\bar{b}'}{}^{-1} h \overline{\bar{b}\bar{b}'} = (\bar{b}\bar{b}'(b,b')^{-1})^{-1} h \bar{b}\bar{b}'(b,b')^{-1}$$
$$= (b,b') h_{b'}{}^{-1} h_b{}^{-1} h h_b h_{b'} (b,b')^{-1}.$$

Therefore, $h_b h_{b'}(b, b')^{-1} h_{bb'}{}^{-1} = 1$ since it belongs to the center of H. Thus $\rho(b, b') = 1_L$.

Remark. This method then applies for example when $H = S_n$ for $n \geqq 3$, $n \neq 6$.

The following gives an example in which the D_b's can not be selected so that $\rho \in K \cdot 1_L$.

Let G be the quaternion group of order 12. G is generated by x and y where $x^3 = y^2$ and $x^6 = y^4 = 1$. G is a non-split extension of the cyclic group $H = [x]$ and a cyclic group $B = [b]$ of order 2 where $\bar{b} = y$, $(b, b) = y^2$, and $G = H \cup yH$.

Let T be the matrix representation of H over a field K of characteristic 2 given by

$$T(x) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Let L be a representation module for T. Then

$$\mathrm{Hom}_{KH}(L, L) = \left\{ \begin{bmatrix} a & c \\ 0 & a \end{bmatrix} \text{ where } a, c \in K \right\}$$

T is indecomposable since $\mathrm{Hom}_{KH}(L, L)$ is completely primary.

Since $T^{(\bar{b})}(x) = T(x)$, D_b can be chosen to be any invertible element of

$\mathrm{Hom}_{KH}(L,L)$. D_1 must be βI_2 for some $\beta\neq 0$, $\in K$ since $\rho(1,1) = D_1D_1(1,1)^{-1}D_1^{-1} = D_1$. Then

$$\rho(b,b) = D_bD_b(b,b)^{-1}D_1^{-1} = \begin{bmatrix}\alpha & \gamma\\ 0 & \alpha\end{bmatrix}^2 \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}\beta^{-1} & 0\\ 0 & \beta^{-1}\end{bmatrix}$$

$$= \alpha^2\beta^{-1}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} \neq I_2$$

for every choice of α, β, and γ. Thus the D_b's can not be selected so that $\rho\in K\cdot 1_L$.

In this example, $S = B$ and L^G is indecomposable. Since B is cyclic of order 2 and the characteristic of K is 2, there exists only the trivial K-factor set for S (see [7]). Thus for the results of § 3 to hold for this group it would be necessary to have $L^G = I(G)\otimes M$ where $I(G) = KB$ with module operation $\#$ given by $\bar{b}x \# s = b\cdot s$ and M is L extended to be a left KG-module. It is easily verified by using matrices that no such M exists.

The construction presented can be used to prove the following theorem which is an extension of the known theorem for the irreducible case ([3], p. 353).

THEOREM 4. *Let $G = G_1\times G_2$ be the direct product of G_1 and G_2. Let M be an indecomposable left KG-module which is (G,G_2) projective. Then, M can be expressed as the outer tensor product $M = L_1 \# L_2$ where L_i is an indecomposable left KG_i-module, $i = 1,2$.*

Proof. Since M is (G,G_2)-projective, it is a component of $(M_{G_2})^G$ (see [3], p. 427). Since M is indecomposable, it is then a component of $(L_2)^G$ for one of the indecomposable KG_2-components L_2 of M_{G_2}.

Then, L_2 is KG_2-isomorphic to $(g_1,g_2)L_2 = (g_1,1)L_2$, for every $(g_1,g_2) \in G_1\times G_2$. G is the split extension of G_2 by G_1. Thus, $S = G_1$. For each $g_1\in G_1$, since $(1,g_2)^{g_1} = (g_1,1)^{-1}(1,g_2)(g_1,1) = (1,g_2)$, D_{g_1} can be taken to be 1_{L_2}. Then, $\rho(g_1,g_1') = 1_{L_2}$, for every $g_1,g_1'\in G_1$.

Therefore, the results of § 3 are available. It follows that $M\cong L_1^*\otimes L_2^*$ where L_1^* is an indecomposable left ideal L_1 of KG_1 which is made into a KG-module by $(g_1,g_2)s = g_1s$ and L_2^* is L_2 made into a KG-module by $(g_1,g_2)l = g_2l$. Since M is indecomposable, L_1^* and L_2^* are indecomposable KG-modules.

Then, M is KG-isomorphic to $L_1 \# L_2$.

Remark. It has been proven by Conlon in [2] that when L is an indecomposable left KH-module, K algebraically closed, the indecomposable

components of L^G are determined by the indecomposable components of a twisted group algebra of S. In the notation here, the components would be determined by $(KS)_{\beta^{-1}}$ where β is obtained by decomposing $\rho(b,b')$ using (1), i.e., $\rho(b,b')=\beta(b,b')\cdot 1_L+n(b,b')$. The method of proof is to use the relationship between the splitting of a left module into submodules and the splitting of its inverse ring of endomorphisms into left ideals. $(KS)_{\beta^{-1}}$ is anti-isomorphic to $\mathrm{Hom}_{KG}(L^G,L^G)$ modulo a nilpotent ideal (and to $\mathrm{Hom}_{KS}(L^S,L^S)$ modulo a nilpotent ideal) and thus determines the splitting of the inverse ring of endomorphisms.

THE INSTITUTE FOR ADVANCED STUDY;
UNIVERSITY OF ILLINOIS, URBANA.

REFERENCES.

[1] A. H. Clifford, "Representations induced in an invariant subgroup," *Annals of Mathematics*, vol. 38 (1937), pp. 533-550.

[2] S. B. Conlon, "Twisted group algebras and their representations," *Journal of the Australian Mathematical Society*, vol. 4 (1964).

[3] C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, New York, 1962.

[4] N. Jacobson, *The Theory of Rings*, New York, 1943.

[5] P. A. Tucker, "On the reduction of induced representations of finite groups," *American Journal of Mathematics*, vol. 84 (1962), pp. 400-421.

[6] ———, "Note on the reduction of induced representations," *ibid.*, vol. 85 (1963), pp. 53-58.

[7] ———, "Endomorphism ring of an induced module," *Michigan Mathematical Journal*, vol. 12 (1965), pp. 197-202.

AFFINE EMBEDDINGS OF COMPLEX ANALYTIC HOMOGENEOUS SPACES.[1]

By G. Hochschild and G. D. Mostow.

Introduction. Let G be a faithfully representable complex analytic group, and let $\boldsymbol{R}(G)$ be the algebra of the complex analytic representative functions on G. In the study of the structure of $\boldsymbol{R}(G)$, an important role is played by certain subalgebras, which we have called basic subalgebras. These subalgebras admit an interpretation as the algebras of the polynomial functions belonging to structures of G as affine algebraic varieties, compatible with the holomorphic structure and such that the right translations effected on G by the elements of G are automorphisms of the variety structure. Thus every complex analytic linear group admits an algebraic variety structure that is stable under translations from one side but, of course, not both sides, in general.

It is the purpose of the present paper to investigate this type of algebraic variety structure more generally on homogeneous spaces $L\backslash G$ of G with respect to closed complex Lie subgroups L. As is already suggested by the known results on homogeneous spaces of algebraic linear groups [1], it turns out that one should then admit variety structures somewhat more general than affine varieties; namely, open subvarieties of affine varieties, called quasi-affine varieties. These structures are intimately linked to the general representation theory of G, and we view them in this context.

The general algebraic-geometric and representation theoretical setting for the present investigation is given in Section 1. Then we proceed, in Section 2, to show that all quasi-affine structures on homogeneous spaces $L\backslash G$ arise as homogeneous spaces of algebraic hulls of G with respect to algebraic subgroups.

In Sections 3 and 4, we deal with the case of reductive subgroups L of G, showing that then $L\backslash G$ always has affine structures, that all quasi-affine structures are necessarily affine and arise naturally from affine structures on G, and that the affine structures on G are precisely the left stable basic subalgebras of $\boldsymbol{R}(G)$ in the sense of [8]. In Section 5, we show that the quasi-affine structures on $L\backslash G$, where L is not necessarily reductive, can be lifted

[1] Research done under NSF Grants GP-1610 and GP-1662.
Received October 1, 1964.

to affine structures on G if and only if L has only a finite number of connected components.

Section 6 is preliminary to Section 7 and is concerned with the separation of the points of a homogeneous space by appropriate subalgebras of $\boldsymbol{R}(G)$. Section 7 gives necessary and sufficient conditions for the existence of quasi-affine structures for a given $L\backslash G$. In Section 8, we show by means of examples that these conditions are independent.

What is taken for granted from algebraic geometry and the general theory of algebraic groups consists only of standard facts. Here, it is appropriate to use the 'concrete' algebraic geometry over a fixed algebraically closed base field (actually, the field of the complex numbers), such as is developed in [2]. One or two results of [3] are involved with regard to the foundations of the theory of algebraic groups. The underlying structure and representation theory of complex Lie groups, on which we rely heavily, is that of [7], [8] and [9].

1. Homogeneous quasi-affine structures. Let V be an irreducible algebraic variety over an algebraically closed field K. We denote by $\boldsymbol{F}(V)$ the field of the K-valued rational functions on V. An everywhere defined rational function will be called a *polynomial function.* The K-valued polynomial functions on V constitute a K-subalgebra $\boldsymbol{P}(V)$ of $\boldsymbol{F}(V)$, which we shall call the *algebra of the polynomial functions on* V. We shall say that V is *quasi-affine* if it is isomorphic with an open subvariety of an affine algebraic variety. This will be the case if and only if there is a finitely generated subalgebra B of $\boldsymbol{P}(V)$ such that the evaluation map of V into the affine variety $\boldsymbol{V}(B)$ of all K-valued specializations of B is an isomorphism of V onto an open subvariety of $\boldsymbol{V}(B)$. We shall then call B an *affine carrier* of V. The field of fractions of every affine carrier of V coincides with $\boldsymbol{F}(V)$. An algebraic variety V is affine if and only if $\boldsymbol{P}(V)$ is finitely generated as a K-algebra and the evaluation map of V into $\boldsymbol{V}(\boldsymbol{P}(V))$ is an isomorphism.

Let G be an abstract group, and suppose that we are given an action, from the right, of G by automorphisms of the algebraic variety V. This is to mean that, to each element x of G, there is attached an automorphism $\zeta \rightarrow \zeta \cdot x$ of V such that, for all elements x and y of G and all points ζ of V, $(\zeta \cdot x) \cdot y = \zeta \cdot (xy)$. Then we obtain a representation of G by algebra automorphisms $f \rightarrow x \cdot f$ of $\boldsymbol{P}(V)$, where $(x \cdot f)(\zeta) = f(\zeta \cdot x)$. Similarly, G acts by field automorphisms on $\boldsymbol{F}(V)$.

Given such an action of a group G on the algebraic variety V, we shall

say that V is *G-homogeneous* if the following two conditions are satisfied: (1) G acts transitively on V; (2) the associated representation of G on $\boldsymbol{P}(V)$ is locally finite, i.e., every element of $\boldsymbol{P}(V)$ lies in a finite-dimensional G-stable K-subspace of $\boldsymbol{P}(V)$.

PROPOSITION 1.1. *Let V be an irreducible algebraic variety over the algebraically closed field K. Suppose that $\boldsymbol{P}(V)$ separates the points of V and that there is given an action of a group G on V such that V is G-homogeneous. Then V is quasi-affine, and every finitely generated G-stable subalgebra of $\boldsymbol{P}(V)$ whose field of fractions coincides with $\boldsymbol{F}(V)$ is an affine carrier of V. Moreover, $\boldsymbol{P}(V)$ contains every finite-dimensional G-stable K-subspace of $\boldsymbol{F}(V)$.*

Proof. Since $\boldsymbol{P}(V)$ separates the points of V, $\boldsymbol{F}(V)$ is a purely inseparable algebraic extension field of the field of fractions of $\boldsymbol{P}(V)$. On the other hand, the set of normal points of V is non-empty so that we conclude from the transitivity of the action of G on V that every point of V is normal. This implies that $\boldsymbol{P}(V)$ is integrally closed in $\boldsymbol{F}(V)$. Together with the first statement of this proof, this shows that the field of fractions of $\boldsymbol{P}(V)$ coincides with $\boldsymbol{F}(V)$. Now $\boldsymbol{F}(V)$ is a finitely generated extension field of K. Hence there is a finitely generated subalgebra of $\boldsymbol{P}(V)$ whose field of fractions is $\boldsymbol{F}(V)$. By condition (2) above, every finitely generated subalgebra of $\boldsymbol{P}(V)$ is contained in a finitely generated G-stable subalgebra of $\boldsymbol{P}(V)$. Thus there exists a finitely generated G-stable subalgebra B of $\boldsymbol{P}(V)$ whose field of fractions is $\boldsymbol{F}(V)$.

Now let us consider the evaluation morphism $\psi\colon V \to \boldsymbol{V}(B)$ defined by $\psi(\zeta)(b) = b(\zeta)$, where $\zeta \in V$ and $b \in B$. The representation of G on B defines an action of G on $\boldsymbol{V}(B)$; for x in G and τ in $\boldsymbol{V}(B)$, the transform $\tau \cdot x$ is given by $(\tau \cdot x)(b) = \tau(x \cdot b)$, for all b in B. The above morphism ψ is evidently a G-morphism: $\psi(\zeta \cdot x) = \psi(\zeta) \cdot x$. Since the elements of B are separated by $\psi(V)$, it is clear that $\psi(V)$ is dense in $\boldsymbol{V}(B)$. By the standard theorem on morphisms of algebraic varieties, $\psi(V)$ therefore contains an open subset of $\boldsymbol{V}(B)$. Since G acts transitively on V and ψ is a G-morphism, it follows that $\psi(V)$ is open in $\boldsymbol{V}(B)$. Since B separates the points of V, ψ is injective and therefore has a set theoretical inverse ϕ: $\psi(V) \to V$. On the other hand, the cohomomorphism of ψ sends $\boldsymbol{F}(\boldsymbol{V}(B))$ onto the field of fractions of B, which is $\boldsymbol{F}(V)$. Hence there is a rational map ρ from $\psi(V)$ to V such that $\psi \circ \rho$ and $\rho \circ \psi$ are identity maps. Thus ϕ coincides with ρ on some open subset D (the domain of ρ) of $\psi(V)$. Since ϕ is a G-map, it follows that ϕ is rational also on every transform $D \cdot x$ of D

by an element x of G. Since G acts transitively on $\psi(V)$, we conclude that ϕ is actually a morphism $\psi(V) \rightarrow V$. Thus ψ is an isomorphism of V onto the open subvariety $\psi(V)$ of $\boldsymbol{V}(B)$. This means that V is quasi-affine and that B is an affine carrier of V.

Now let f be an element of a finite-dimensional G-stable K-subspace of $\boldsymbol{F}(V)$. Then there are elements $x_1, \cdots, x_n$ in G such that each transform $x \cdot f$ is a K-linear combination of $x_1 \cdot f, \cdots, x_n \cdot f$. On the other hand, there is a non-zero element p in $\boldsymbol{P}(V)$ such that $pf \in \boldsymbol{P}(V)$, because the field of fractions of $\boldsymbol{P}(V)$ coincides with $\boldsymbol{F}(V)$. Put $u = (x_1 \cdot p) \cdots (x_n \cdot p)$. Then we have $u(x \cdot f) \in \boldsymbol{P}(V)$, for every element x of G. Choose ζ_1 in V such that $u(\zeta_1) \neq 0$. If ζ is an arbitrary point of V there is an element x in G such that $\zeta_1 \cdot x = \zeta$. Now we have $u(x \cdot f) \in \boldsymbol{P}(V)$, whence also $(x^{-1} \cdot u) f \in \boldsymbol{P}(V)$. Since $(x^{-1} \cdot u)(\zeta) = u(\zeta_1) \neq 0$ and $x^{-1} \cdot u \in \boldsymbol{P}(V)$, this shows that f is defined at ζ. Thus $f \in \boldsymbol{P}(V)$, and Proposition 1.1 is proved.

Note. If K is of characteristic 0 then every subfield of $\boldsymbol{F}(V)$ that separates the points of V necessarily coincides with $\boldsymbol{F}(V)$. Hence, in the situation of Proposition 1.1, *if K is of characteristic 0 then every finitely generated G-stable subalgebra of $\boldsymbol{P}(V)$ that separates the points of V is an affine carrier of V.*

We shall be concerned with irreducible algebraic varieties V over the field C of the complex numbers. Suppose we are given an action of a complex analytic group G by automorphisms on such a variety V satisfying the following conditions:

(1) G acts transitively (from the right) on V;

(2) for every polynomial function f on V and every point ζ of V, the map $f_\zeta \colon G \rightarrow C$ defined by $f_\zeta(x) = f(\zeta \cdot x)$ is a holomorphic function on G.

We shall refer to such a structure by saying that V is a *G-variety*. The following lemma will show that a G-variety is always G-homogeneous.

LEMMA 1.2. *Let G be a complex analytic group, and let f be a holomorphic function on G such that the C-space spanned by the translates $x \cdot f$, with x ranging over G, is of countable dimension. Then f is a representative function on G, i.e., the C-space spanned by the translates $x \cdot f$ is actually finite-dimensional.*

Proof. Let $(f_i)_{i=1,2,\cdots}$ be a C-basis for the space of the translates $x \cdot f$. For each x in G, write $x \cdot f = \sum_i c_i(x) f_i$, and let $n(x)$ denote the largest

index i such that $c_i(x) \neq 0$. For each positive integer k, let G_k denote the closure in G of the set of all x in G such that $n(x) = k$. Then the union of these sets G_k is G. Since G is locally compact, it follows that at least one of these sets G_k contains a non-empty open subset V of G. Thus $n(x) = k$ for all points x of some dense subset D of V, so that for each x in D we have $x \cdot f = \sum_{i=1}^{k} c_i(x) f_i$. We can find elements $x_1, \cdots, x_k$ in G such that the determinant with the entries $f_i(x_j)$ is different from 0. Hence we can solve the relations $f(x_j x) = \sum_{i=1}^{k} c_i(x) f_i(x_j)$ for the coefficients $c_i(x)$, obtaining $c_i(x) = \sum_{j=1}^{k} c_{ji} f(x_j x)$, for each x in D, where $c_{ji} \in C$. Put $d_i(y) = \sum_{j=1}^{k} c_{ji} f(x_j y)$, for all y in G. Then each d_i is a holomorphic function on G, and $x \cdot f = \sum_{i=1}^{k} d_i(x) f_i$ for each x in D. For each y in G, define the holomorphic function g_y on G by $g_y(x) = f(yx) - \sum_{i=1}^{k} d_i(x) f_i(y)$. Then g_y vanishes on D, and hence also on V. Since G is connected, the vanishing of the holomorphic function g_y on the non-empty open set V implies that $g_y = 0$. This, for all elements y of G, means that $x \cdot f = \sum_{i=1}^{k} d_i(x) f_i$ for all elements x of G, so that Lemma 1.2 is proved.

In applying Lemma 1.2 later on, we shall make use of the fact that *if V is any irreducible algebraic variety over an algebraically closed field K then $\boldsymbol{P}(V)$ is of countable dimension as a vector space over K.* This is seen as follows. Let W be a non-empty open affine subvariety of V. Then $\boldsymbol{P}(W)$ is a finitely generated K-algebra and is therefore of countable dimension as a vector space over K. Since the restriction map $\boldsymbol{P}(V) \to \boldsymbol{P}(W)$ is injective, it follows that $\boldsymbol{P}(V)$ is also of countable dimension as a vector space over K.

If G is a complex analytic group we denote by $\boldsymbol{R}(G)$ the C-algebra of all complex analytic representative functions on G and by $\boldsymbol{H}(G)$ the C-algebra of all holomorphic functions on G. We shall denote by $\boldsymbol{M}(G)$ the field of fractions of the integral domain $\boldsymbol{H}(G)$. We regard $\boldsymbol{H}(G)$ as a left G-module, with $(x \cdot f)(y) = f(yx)$. Then $\boldsymbol{R}(G)$ is precisely the maximum locally finite G-submodule of $\boldsymbol{H}(G)$. The action of G on $\boldsymbol{H}(G)$ extends to an action of G by field automorphisms on $\boldsymbol{M}(G)$. One shows exactly as in the last part of the proof of Proposition 1.1 that $\boldsymbol{R}(G)$ is also the maximum locally finite G-submodule of $\boldsymbol{M}(G)$.

If V is a G-variety and ζ is a point of V we denote by ζ^* the C-algebra monomorphism of $\boldsymbol{P}(V)$ into $\boldsymbol{H}(G)$ that is given by $\zeta^*(f) = f_\zeta$. Clearly, ζ^* is also a G-module monomorphism.

THEOREM 1.3. *Let G be a complex analytic group, V a G-variety, ζ a point of V. Then V is G-homogeneous and $\zeta^*(\boldsymbol{P}(V)) \subset \boldsymbol{R}(G)$. If $\boldsymbol{P}(V)$ separates the points of V then V is quasi-affine, every finitely generated G-stable separating subalgebra B of $\boldsymbol{P}(V)$ is an affine carrier of V, and $\zeta^*(\boldsymbol{P}(V))$ coincides with the maximum locally finite G-submodule of the field of fractions of $\zeta^*(B)$.*

Proof. We known from the remark following the proof of Lemma 1.2 that $\boldsymbol{P}(V)$ has a countable G-basis. Since ζ^* is a G-homomorphism, we may therefore apply Lemma 1.2 to conclude that $\zeta^*(\boldsymbol{P}(V)) \subset \boldsymbol{R}(G)$. Since ζ^* is injective, this implies that the representation of G on $\boldsymbol{P}(V)$ is locally finite, so that V is G-homogeneous.

Now suppose that $\boldsymbol{P}(V)$ separates the points of V. Then it is clear that the isotropy group, L say, of ζ in G is precisely the group consisting of all elements x in G such that $f \cdot x = f$ for every element f of $\zeta^*(\boldsymbol{P}(V))$, were $f \cdot x$ is the element of $\boldsymbol{R}(G)$ that is defined by $(f \cdot x)(y) = f(xy)$. Since the anti-representation of G by right translations on $\boldsymbol{R}(G)$ is locally finite and complex analytic, this shows, incidentally, that *L is a closed complex Lie subgroup of G.*

By Proposition 1.1 and the note following its proof, we have that V is quasi-affine and that every finitely generated G-stable separating subalgebra B of $\boldsymbol{P}(V)$ is an affine carrier of V.

Now let M denote the maximum locally finite G-submodule of the field of fractions of $\zeta^*(B)$. Since the field of fractions of B coincides with $\boldsymbol{F}(V)$, it contains $\boldsymbol{P}(V)$, so that the field of fractions of $\zeta^*(B)$ contains $\zeta^*(\boldsymbol{P}(V))$. Since $\zeta^*(\boldsymbol{P}(V)) \subset \boldsymbol{R}(G)$, it is locally finite, and our last result therefore gives that $\zeta^*(\boldsymbol{P}(V)) \subset M$. Now let f be any element of M. Then there is a non-zero element g of $\zeta^*(B)$ such that $gf \in \zeta^*(B)$. Now we observe that ζ^* extends canonically to a field monomorphism of $\boldsymbol{F}(V)$ into $\boldsymbol{M}(G)$, and that this field monomorphism is also a G-module monomorphism. Denoting this extension of ζ^* still by ζ^*, we have from the above that $f = \zeta^*(f^*)$, where f^* is an element of the field of fractions of B, i.e., where $f^* \in \boldsymbol{F}(V)$. Since the G-module generated by f is finite-dimensional and ζ^* is a G-module monomorphism, we conclude that f^* lies in a finite-dimensional G-submodule of $\boldsymbol{F}(V)$. By Proposition 1.1, we have therefore $f^* \in \boldsymbol{P}(V)$, so that $f \in \zeta^*(\boldsymbol{P}(V))$. Thus $M \subset \zeta^*(\boldsymbol{P}(V))$, and our proof of Theorem 1.3 is complete.

Note. The last argument, with $\boldsymbol{P}(V)$ in the place of B, shows that *the intersection with $\boldsymbol{R}(G)$ of the field of fractions of $\zeta^*(\boldsymbol{P}(V))$ coincides with $\zeta^*(\boldsymbol{P}(V))$.*

In the situation of Theorem 1.3, with $\boldsymbol{P}(V)$ separating the points of V, if L is the isotropy subgroup of ζ in G we may identify the set V with the set $L\backslash G$ of the cosets Lx, and we may describe the situation by saying that $L\backslash G$ is endowed with the structure of a quasi-affine G-homogeneous G-variety, the action of G being the canonical one and the algebraic variety structure being compatible with the canonical structure of $L\backslash G$ as a complex analytic variety in the sense that the polynomial functions are holomorphic. In the sequel, when we say that $L\backslash G$ is endowed with a *quasi-affine structure,* this is to imply that the structure is of the kind just described. If V is actually affine we shall say that this structure on $L\backslash G$ is an *affine structure.* If B is an affine carrier of V then we shall call $\zeta^*(B)$ *an affine carrier of* $L\backslash G$ *in* $\boldsymbol{R}(G)$.

It will be convenient to view a quasi-affine structure on $L\backslash G$ simply as the subalgebra $\zeta^*(\boldsymbol{P}(V))$ of $\boldsymbol{R}(G)$. In view of what we have already proved, the following definition is equivalent to the above.

Definition 1.4. Let G be a complex analytic group, L a closed complex Lie subgroup of G. A quasi-affine structure for $L\backslash G$ in $\boldsymbol{R}(G)$ is a subalgebra P of $\boldsymbol{R}(G)$ satisfying the following conditions:

(1) P is left G-stable, and its elements are fixed under the action of L on $\boldsymbol{R}(G)$ from the right.

(2) There is a finitely generated left G-stable subalgebra B of P whose field of fractions coincides with that of P such that the evaluation map induces a bijection of $L\backslash G$ onto an open subvariety of $\boldsymbol{V}(B)$.

(3) The intersection with $\boldsymbol{R}(G)$ of the field of fractions of P coincides with P.

It may be appropriate to recall that the algebra B of (2) is an affine carrier for $L\backslash G$ in $\boldsymbol{R}(G)$, and that, in fact, every finitely generated left G-stable subalgebra of P that separates the points of $L\backslash G$ satisfies (2) and is an affine carrier for $L\backslash G$ in $\boldsymbol{R}(G)$.

2. Quasi-affine structures and algebraic group hulls. *From now on, it will be assumed throughout that the complex analytic group G under consideration has a faithful finite-dimensional complex analytic representation or, equivalently, that $\boldsymbol{R}(G)$ separates the points of G.* We denote by A the group of the *proper automorphisms* of $\boldsymbol{R}(G)$, i.e., of all algebra automorphisms of $\boldsymbol{R}(G)$ that commute with every right translation $f \rightarrow f \cdot x$ effected by an element x of G on $\boldsymbol{R}(G)$. The map that associates with every

element x of G the left translation $f \rightarrow x \cdot f$ effected by x on $\boldsymbol{R}(G)$ is a group monomorphism $G \rightarrow A$ through which we identify G with a subgroup of A. Every left G-stable subspace S of $\boldsymbol{R}(G)$ is also A-stable, and we shall denote by A_S or G_S the restriction image of A or G in the group of all linear automorphisms of S.

A subset of $\boldsymbol{R}(G)$ is said to be *fully stable* if it is stable under the right and left G-translations as well as under the involution $f \rightarrow f'$, where $f'(x) = f(x^{-1})$. If S is a finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$ then A_S is the group of all proper automorphisms of S and has a natural structure of an irreducible affine algebraic group with S as the algebra of all polynomial functions, where an element f of S is regarded as a C-valued function on A_S by $f(\alpha) = \alpha(f)(1)$. This structure of A_S defines also the structure of a complex analytic group on A_S. The subgroup G_S of A_S is a complex analytic subgroup of A_S, and the canonical map $G \rightarrow G_S$ is an epimorphism of complex analytic groups. Moreover, G_S is algebraically dense in A_S and contains the commutator subgroup of A_S. In fact, G contains the commutator subgroup of A. If T is a finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$ such that $T \subset S$ then the restriction map $A_S \rightarrow A_T$ is an epimorphism of algebraic groups, and A is the projective limit of this system of algebraic group epimorphisms. Thus we may view A as a *pro-algebraic* group.

If Y is a left G-submodule of S such that S is the smallest fully stable subalgebra of $\boldsymbol{R}(G)$ containing Y then the restriction map $A_S \rightarrow A_Y$ is a group isomorphism, so that A_Y is also an irreducible affine algebraic group and a complex analytic group, isomorphic in each sense with A_S, in a natural way. In particular, if Y is finite -dimensional, A_Y is thus an irreducible algebraic subgroup of the full linear group on Y, and so is an irreducible linear algebraic group.

If X is a subset of A we shall denote by $\boldsymbol{R}(G)^X$ the algebra of all X-fixed elements of $\boldsymbol{R}(G)$. If Y is a subset of $\boldsymbol{R}(G)$ then A^Y denotes the *fixer* of Y in A, i. e., the subgroup consisting of all elements of A that leave the elements of Y fixed. We shall use the same notation in connection with A_S. Finally, we denote by Y' the image of Y under the involution $f \rightarrow f'$ of $\boldsymbol{R}(G)$.

LEMMA 2.1. *Let S be a finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$, B a finitely generated left G-stable subalgebra of S. Then the elements of B are constant on the cosets $(A_S)^{B'}\alpha$ of $(A_S)^{B'}$ in A_S and hence may be regarded as polynomial functions on the algebraic variety $(A_S)^{B'}\backslash A_S$. This algebraic variety is quasi-affine, and B is an affine carrier for it in $\boldsymbol{R}(G)$.*

Proof. In proving the first statement, we use the fact that if $\beta \in A$ and $f \in R(G)$ then $\beta(f)(1) = \beta^{-1}(f')(1)$. This is evident for β in G and extends to A, because G_S is algebraically dense in A_S for every S. Now let β be an element of $(A_S)^{B'}$ and α an element of A_S. Then we have, for every element b of B,

$$\begin{aligned} b(\beta\alpha) = (\beta\alpha)(b)(1) = \beta(\alpha(b))(1) = \beta^{-1}(\alpha(b)')(1) \\ = \alpha(b)'(1) = \alpha(b)(1) = b(\alpha), \end{aligned}$$

which establishes the first assertion of our lemma.

Next we show that B separates the points of $(A_S)^{B'}\backslash A_S$. Suppose that α and β are elements of A_S such that $b(\alpha) = b(\beta)$ for every element b of B. We must show that then α^{-1} and β^{-1} coincide on B'. Now it is known that the inverse of an element γ of A is given by $\gamma^{-1}(f)(x) = (x^{-1}\cdot f')(\gamma)$. In fact, this is evident for γ in G and extends to A, because G_S is algebraically dense in A_S for every S. Using this, we obtain, for every element b of B, and every element x of G,

$$\beta^{-1}(b')(x) = (x^{-1}\cdot b)(\beta) = (x^{-1}\cdot b)(\alpha) = \alpha^{-1}(b')(x),$$

so that indeed $\beta^{-1}(b') = \alpha^{-1}(b')$.

The polynomial functions on the algebraic variety $(A_S)^{B'}\backslash A_S$ may be identified with the elements of S that are constant on the cosets $(A_S)^{B'}\alpha$. Hence it is clear that the variety $(A_S)^{B'}\backslash A_S$ is A_S-homogeneous for the canonical right action of A_S. Hence we may apply Proposition 1.1 and the note at the end of its proof to conclude that $(A_S)^{B'}\backslash A_S$ is quasi-affine and that B is an affine carrier. This completes the proof of Lemma 2.1.

THEOREM 2.2. *Let L be a closed complex Lie subgroup of G. Suppose that there is a quasi-affine structure P for $L\backslash G$ in $\mathbf{R}(G)$. Let S be a fully stable finitely generated subalgebra of $\mathbf{R}(G)$ containing a left G-stable affine carrier B for $L\backslash G$. Then $(A^{P'})_S$ coincides with the algebraic subgroup $(A_S)^{B'}$ of A_S. The canonical homomorphism $G \to A_S$ induces an isomorphism of G-varieties and complex analytic manifolds $L\backslash G \to (A^{P'})_S\backslash A_S$. In particular, $(A^{P'})_S G_S = A_S$* and *$(A^{P'})_S \cap G_S = L_S$. Moreover, $A^{P'} = A$, $A^{P'} \cap G = L$* and *$P' = \mathbf{R}(G)^{A^{P'}}$.*

Proof. Evidently, $(A^{P'})_S \subset (A_S)^{B'}$. On the other hand, we know that the intersection with $\mathbf{R}(G)$ of the field of fractions of B coincides with P. Hence also the intersection with $\mathbf{R}(G)$ of the field of fractions of B' coincides with P'. Hence, if an element of A leaves the elements of B' fixed, it also

leaves the elements of P' fixed. It follows that $(A_S)^{B'} \subset (A^{P'})_S$, and the first assertion of our theorem is proved.

By Lemma 2.1, the evaluation map induces an isomorphism of the algebraic variety $(A_S)^{B'}\backslash A_S$ onto an open subvariety W of $\boldsymbol{V}(B)$. Clearly, $L_S \subset (A_S)^{B'}$, so that the canonical homomorphism $G \to A_S$ induces a morphism of complex analytic varieties $L\backslash G \to (A_S)^{B'}\backslash A_S$. On the other hand, since B is an affine carrier for $L\backslash G$ in $\boldsymbol{R}(G)$, the evaluation map induces an isomorphism of the algebraic variety $L\backslash G$ onto an open subvariety W_1 of $\boldsymbol{V}(B)$. These maps evidently form a commutative diagram

$$\begin{tikzcd} (A_S)^{B'}\backslash A_S \arrow[dr] & \\ & \boldsymbol{V}(B) \\ L\backslash G \arrow[uu] \arrow[ur] & \end{tikzcd}$$

By the commutativity of the diagram, we have $W_1 \subset W$. The image of $L\backslash G$ in $(A_S)^{B'}\backslash A_S$ evidently coincides with the canonical image of the subgroup $(A_S)^{B'}G_S$ of A_S. Since W_1 is an open algebraic subvariety of W, its inverse image in $(A_S)^{B'}\backslash A_S$ is an open algebraic subvariety, and therefore is also topologically open. Hence $(A_S)^{B'}G_S$ is an open subgroup of A_S, whence $(A_S)^{B'}G_S = A_S$ and $W_1 = W$. Now it is clear from the commutavity of the diagram and the fact that the slanted maps are isomorphisms onto W that the vertical map $L\backslash G \to (A_S)^{B'}\backslash A_S$ is an isomorphism of G-varieties and complex analytic manifolds. Evidently, this implies that $(A_S)^{B'} \cap G = L_S$.

Clearly, the kernel of the restriction epimorphism $A \to A_S$ is contained in $A^{B'} = A^{P'}$. Since $(A^{P'})_S G_S = A_S$, we have therefore also $A^{P'}G = A$. We may take S such that the canonical map $G \to G_S$ is injective. Since $(A^{P'})_S \cap G_S = L_S$, we have therefore $A^{P'} \cap G = L$.

Finally, since B is an affine carrier of $(A^{P'})_S\backslash A_S$, the algebra of the polynomial functions on $(A^{P'})_S\backslash A_S$ is contained in the field of fractions of B. Hence $(S^{A^{P'}})'$ is contained in the field of fractions of B, and therefore lies in P. Hence we have $S^{A^{P'}} \subset P'$. Since we may take S so as to contain any given element of $\boldsymbol{R}(G)$, this shows that $\boldsymbol{R}(G)^{A^{P'}} \subset P'$. Our proof of Theorem 2.2 is now complete.

There is a partial converse of Theorem 2.2, giving a characterization of quasi-affine structures in $\boldsymbol{R}(G)$, as follows.

THEOREM 2.3. *Let P be a subalgebra of $\boldsymbol{R}(G)$ satisfying the following conditions:*

(1) *P is left G-stable;* (2) *the field of fractions of P is a finitely*

generated extension field of C and its intersection with $\boldsymbol{R}(G)$ coincides with P; (3) $A^{P'}G = A$.

Then P is a quasi-affine structure for $(A^{P'} \cap G)\backslash G$ in $\boldsymbol{R}(G)$.

Proof. By (1) and the first part of (2), there is a finitely generated left G-stable subalgebra B of P whose field of fractions coincides with that of P. There is a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ such that $B \subset S$. By Lemma 2.1, the algebraic variety $(A_S)^{B'}\backslash A_S$ is quasi-affine and has B as an affine carrier. From the second part of condition (2), we have that $(A_S)^{B'} = (A^{P'})_S$. By condition (3), we have $(A^{P'})_S G_S = A_S$. Hence the injection $G_S \rightarrow A_S$ induces a bijection

$$((A^{P'})_S \cap G_S)\backslash G_S \rightarrow (A^{P'})_S\backslash A_S.$$

Since $(A^{P'})_S = (A_S)^{B'}$ and $B' \subset S$, it is clear that $(A^{P'})_S \cap G_S = (A^{P'} \cap G)_S$. Taking S large enough for the canonical map $G \rightarrow G_S$ to be injective, we have from the above that the canonical map $G \rightarrow A_S$ induces a complex analytic bijection $(A^{P'} \cap G)\backslash G \rightarrow (A^{P'})_S\backslash A_S$. Hence the quasi-affine structure on $(A^{P'})_S\backslash A_S$ can be transported into one of $(A^{P'} \cap G)\backslash G$, for which B is an affine carrier. As a subalgebra of $\boldsymbol{R}(G)$, this quasi-affine structure of $(A^{P'} \cap G)\backslash G$ is therefore P.

3. **Existence of affine structures.** We must recall some definitions and results concerning the structure of $\boldsymbol{R}(G)$. Our assumption that $\boldsymbol{R}(G)$ separates the points of G is still in force. The additive group of all complex analytic homomorphisms of G into C is denoted by $\mathrm{Hom}(G, C)$. The composites of the elements of $\mathrm{Hom}(G, C)$ with the exponential map of the additive group C into the multiplicative group C^* of the non-zero complex numbers make up a subgroup Q of the multiplicative group of all complex analytic homomorphisms of G into C^*. It is known from [8] that there are left G-stable finitely generated subalgebras B of $R(G)$ such that Q is free over B and $\boldsymbol{R}(G) = B[Q]$. Such a subalgebra B is called a *left stable basic subalgebra* of $\boldsymbol{R}(G)$. It necessarily contains $\mathrm{Hom}(G, C)$.

Let $\tau \colon B \rightarrow C$ be a specialization of a left stable basic subalgebra B of $\boldsymbol{R}(G)$. Define the homomorphism $\sigma \colon B' \rightarrow \boldsymbol{R}(G)$ as follows: for f in B' and x in G, $\sigma(f)(x) = \tau(x^{-1} \cdot f')$. One verifies directly that σ commutes with the right G-action, i. e., that $\sigma(f \cdot y) = \sigma(f) \cdot y$, for all f in B' and all y in G. Since Q is free over B and $B[Q] = \boldsymbol{R}(G)$, we have also (since $Q = Q'$) that Q is free over B' and $B'[Q] = \boldsymbol{R}(G)$. Now B', like B, contains $\mathrm{Hom}(G, C)$. Since σ commutes with the right G-action and leaves the constants fixed we have $\sigma(h) = h + \sigma(h)(1) = h + \tau(h')$, for every element h of $\mathrm{Hom}(G, C)$.

Every element q of Q is of the form $\exp(h_q)$, with a uniquely determined element h_q of $\mathrm{Hom}(G, C)$. We extend σ to an algebra endomorphism ρ of $\boldsymbol{R}(G)$ such that $\rho(q) = \exp(\tau(h_q'))q$, for every element q of Q. Now ρ is an algebra and right G-module endomorphism of $\boldsymbol{R}(G)$. It is known from [6] that this implies that ρ is actually an algebra *automorphism* of $\boldsymbol{R}(G)$ and thus an element of A. Moreover, from the definition of ρ, we have $\rho(\exp(h)) = \exp(\rho(h)(1))\exp(h)$, for every element h of $\mathrm{Hom}(G, C)$. Now, by [7, Th. 5.1], this is precisely the property that characterizes G as a subgroup of A. Thus there is an element x in G such that $\rho(f) = x \cdot f$, for every element f of $\boldsymbol{R}(G)$. In particular, this gives $\tau(b) = b(x^{-1})$, for every element b of B. Thus every specialization of B is the evaluation at an element of G.

Moreover, since Q is free over B and $\boldsymbol{R}(G) = B[Q]$, it is clear that the intersection with $\boldsymbol{R}(G)$ of the field of fractions of B coincides with B, and that B separates the points of G. Thus we have the following result.

PROPOSITION 3.1. *Every left stable basic subalgebra of $\boldsymbol{R}(G)$ is an affine structure for G in $\boldsymbol{R}(G)$.*

It is known from [8] that there are actually left stable basic subalgebras B having the further stability property that the subalgebra B_s consisting of the *semisimple* representative functions in B is both left and right G-stable. Such a subalgebra is called a *normal basic subalgebra.* A *regular subalgebra* of $\boldsymbol{R}(G)$ is a finitely generated fully stable subalgebra containing a normal basic subalgebra of $\boldsymbol{R}(G)$. The position of such a subalgebra, S say, in $\boldsymbol{R}(G)$, and the position of G_S in A_S are particularly transparent, and the known results in this connection lead to a short proof of the existence theorem we give below, as Theorem 3.2. Before stating this, we recall (from [7] and [10]) some basic facts concerning reductive groups.

A complex Lie group H is called *reductive* if it satisfies the following conditions: (1) H has only a finite number of connected components; (2) H has a faithful finite-dimensional complex analytic representation; (3) every finite-dimensional complex analytic representation of H is semisimple. If a reductive group H is contained as a complex Lie subgroup in a complex analytic group G as above then H is necessarily closed in G. We shall then say simply that H is a reductive subgroup of G.

By a *full* compact subgroup of a complex Lie group L we shall mean a compact subgroup K such that the real Lie algebra of K spans the Lie algebra of L over C and $L = KL_1$, where L_1 denotes the connected component of the identity in L. Every reductive complex Lie group has a full compact subgroup. Conversely, if H is a complex Lie group having a faithful complex analytic representation and a full compact subgroup then H is reductive.

Every reductive complex Lie group has one and only one structure of an affine algebraic group that is compatible with its structure as a complex Lie group, and every complex analytic representation is a rational representation. In particular, the image of a reductive complex Lie group under a finite-dimensional complex analytic representation is always a fully reducible algebraic subgroup of the corresponding full linear group. Conversely, every fully reducible algebraic linear group over C is a reductive complex Lie group.

THEOREM 3.2. *Let L be a reductive subgroup of G. Then G and $L\backslash G$ can be equipped with affine structures such that the canonical map $G \to L\backslash G$ is a morphism of affine algebraic varieties.*

Proof. We choose a regular subalgebra S of $\boldsymbol{R}(G)$ containing a normal basic subalgebra B. By Proposition 3.1, B is an affine structure for G in $\boldsymbol{R}(G)$. Let M be a maximal fully reducible subgroup of A_S such that $L_S \subset M$. By [9, Th. 3.1], $M \cap G_S$ is an algebraic subgroup of M and, as an algebraic group, M is the direct product $(A_S)^{B'} \times (M \cap G_S)$. By Theorem 2.2, the canonical map $G \to (A_S)^{B'}\backslash A_S$ is an isomorphism of algebraic varieties. Moreover, by [9, Th. 3.1], $(A_S)^{B'}$ is isomorphic with the direct product of a finite number of copies of C^* so that, in particular, it is a fully reducible algebraic subgroup of A_S. The subgroup of A_S that is generated by $(A_S)^{B'}$ and L_S is the direct product $(A_S)^{B'} \times L_S$. It is evidently a fully reducible algebraic subgroup of A_S, because both $(A_S)^{B'}$ and L_S are fully reducible. This is known to imply that the corresponding homogeneous space $((A_S)^{B'} \times L_S)\backslash A_S$ is an *affine* algebraic variety; [5, Th. 5.1]. When S is regarded as the algebra of all polynomial functions on A_S, the algebra of all polynomial functions on $((A_S)^{B'} \times L_S)\backslash A_S$ is the subalgebra of S consisting of the elements of S that are constant on the cosets $((A_S)^{B'} \times L_S)\alpha$. We know already that the elements of S that are constant on the coset $(A_S)^{B'}\alpha$ are precisely the elements of B. Hence the algebra of the polynomial functions on $((A_S)^{B'} \times L_S)\backslash A_S$ is the subalgebra ${}^LB$ of B consisting of the elements b in B such that $b \cdot x = b$, for every element x of L. Since the canonical map $G \to (A_S)^{B'}\backslash A_S$ is an isomorphism, it is clear that the canonical map $G \to ((A_S)^{B'} \times L_S)\backslash A_S$ induces a bijection $L\backslash G \to ((A_S)^{B'} \times L_S)\backslash A_S$. This bijection transports the variety structure of $((A_S)^{B'} \times L_S)\backslash A_S$ into the structure of an affine algebraic variety on $L\backslash G$ such that the corresponding affine structure in $\boldsymbol{R}(G)$ is ${}^LB$. Since ${}^LB \subset B$, it is clear that the canonical map $G \to L\backslash G$ is a morphism of algebraic varieties, so that our proof of Theorem 3.2 is complete.

For the proof of the next result we require the following lemma.

LEMMA 3.3. *Let L be a linear complex analytic group, and let L^* be the algebraic group hull of L in the corresponding full linear group. Let M be a maximal fully reducible subgroup of L^*. Then $L^* = ML$.*

Proof. Let U be the maximum unipotent normal subgroup of L^*. Then, by the standard decomposition theorem for algebraic linear groups, L^* is the semidirect product $M \cdot U$. Let ϕ denote the corresponding projection map $L^* \to U$. Since L is normal in L^*, it is clear that ML is a complex analytic subgroup of L^*. Evidently, $\phi(ML) = (ML) \cap U$. Thus $(ML) \cap U$ is connected and is therefore a complex analytic subgroup of the unipotent algebraic linear group U. Hence $(ML) \cap U$ is an algebraic subgroup of U, and so of L^*. Since $(ML) \cap U$ is normal in L^*, it is therefore clear that $M((ML) \cap U)$ is an algebraic subgroup of L^*. Since it contains L, it therefore coincides with L^*. A fortiori, we have $L^* = ML$, and our lemma is proved.

THEOREM 3.4. *Let L be a reductive subgroup of G. Then every quasi-affine structure of $L \backslash G$ is actually affine.*

Proof. Let S be a finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$ containing an affine carrier B for the given quasi-affine structure on $L \backslash G$. Let us write F for $(A_S)^{B'}$. By Theorem 2.2, the canonical homomorphism $G \to A_S$ induces an isomorphism of algebraic varieties $L \backslash G \to F \backslash A_S$. Hence it suffices to show that F is a fully reducible algebraic group, because it is known that $F \backslash A_S$ is then affine. This amounts to showing that F has a full compact subgroup.

Since A_S is an algebraic group hull of G_S, we see from Lemma 3.3 that A_S/G_S has a full compact subgroup. We have $A_S = FG_S$ and $F \cap G_S = L_S$. Hence the complex analytic group A_S/G_S is isomorphic with F/L_S. Hence F/L_S has a full compact subgroup, X say. Since $L_S/(L_S)_1$ is finite, the complete inverse image, Y say, of X in $F/(L_S)_1$ is compact. Since F is an algebraic group, it has only a finite number of connected components, and it follows from the theory of maximal compact subgroups of real Lie groups with finite component groups that the canonical map $F \to F/(L_S)_1$ sends every maximal compact subgroup of F onto a group containing a conjugate of Y. Hence the canonical map $F \to F/L_S$ sends every maximal compact subgroup of F onto a group containing a conjugate of X. Let Z be a maximal compact subgroup of F containing a full compact subgroup of the reductive group L_S. Let T be the smallest complex Lie subgroup of F containing Z. Then the image of T in F/L_S must coincide with F/L_S, and T must contain

L_S. Hence $T = F$, so that Z is a full compact subgroup of F. This completes the proof of Theorem 3.4.

4. Affine structures and basic subalgebras. We shall say that a subgroup X of the group A of proper automorphisms of $\boldsymbol{R}(G)$ is a *pro-algebraic subgroup* if it satisfies the following conditions: (1) for every finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$, X_S is an algebraic subgroup of A_S; (2) X coincides with the canonical image in A of the projective limit of the system of restriction epimorphisms $X_T \to X_S$, where S and T range over the finitely generated fully stable subalgebras of $\boldsymbol{R}(G)$ such that $S \subset T$.

THEOREM 4.1. *Suppose that X is a pro-algebraic subgroup of A such that each element of X is a semisimple automorphism of $\boldsymbol{R}(G)$ and A is the semidirect product $X \cdot G$. Then $\boldsymbol{R}(G)^X$ is a basic subalgebra of $\boldsymbol{R}(G)$, and $A^{\boldsymbol{R}(G)^X} = X$.*

Proof. Note first that, since A is the semidirect product $X \cdot G$ and since the commutator subgroup of A lies in G, the group X is necessarily abelian. Let S be any regular subalgebra of $\boldsymbol{R}(G)$. Then X_S is an abelian algebraic subgroup of A_S consisting entirely of semisimple elements, so that X_S is a fully reducible algebraic subgroup of A_S. Let M be a maximal fully reducible subgroup of A_S such that $X_S \subset M$. It is known from [9, Th. 3.1] and its proof that S contains a normal basic subalgebra B of $\boldsymbol{R}(G)$ such that $M = (A_S)^{B'} \times (M \cap G_S)$, and that $M \cap G_S$ is a maximal reductive subgroup of G_S. Thus $M \cap G_S = H_S$, where H is a maximal reductive subgroup of G.

Now consider the finitely generated fully stable subalgebras T of $\boldsymbol{R}(G)$ such that $S \subset T$. Let $H(T)$ be the subgroup of H consisting of the elements h in H such that h_T belongs to X_T. Since T separates the points of G, the canonical map $H \to H_T$ is an isomorphism. It sends $H(T)$ onto $X_T \cap H_T$. Hence each $H(T)$ is an algebraic subgroup of the algebraic linear group H. Clearly, if $T_1 \subset T_2$ then $H(T_2) \subset H\ T_1)$. Since X is a pro-algebraic subgroup of A and $X \cap G = (1)$, it is clear that $\bigcap_T H(T) = (1)$. Since all the $H(T)$'s are algebraic subgroups of H, this evidently implies that $H(T) = (1)$ for some T, and hence also for every larger T. Hence we may choose the above regular subalgebra S such that $X_S \cap H_S = (1)$.

Since $M = X_S(M \cap G_S)$, it is therefore the semidirect product $X_S \cdot H_S$. Hence, if ρ denotes the restriction to X_S of the projection of M onto its direct factor $(A_S)^{B'}$, according to the decomposition $M = (A_S)^{B'} \times H_S$, then ρ is an algebraic group *isomorphism* $X_S \to (A_S)^{B'}$.

As before, let $Q = \exp(\mathrm{Hom}(G, C))$. We know from [9, proof of Th.

3.1] that $(A_S)^{B'}$ is isomorphic with $\mathrm{Hom}(Q \cap S, C^*)$ by the map $\alpha \to \alpha^*$, where $\alpha^*(q) = \alpha(q)q^{-1}$, for every element q of Q. Moreover, $Q \cap S$ is a finitely generated free abelian group, and if $(q_1, \cdots, q_n)$ is a free basis for this group then the map sending each α onto the n-tuple $(\alpha^*(q_1), \cdots, \alpha^*(q_n))$ of non-zero complex numbers is an algebraic group isomorphism $(A_S)^{B'} \to (C^*)^n$. The group $\mathrm{Hom}((C^*)^n, C^*)$ of the complex analytic, i.e., rational homomorphisms $(C^*)^n \to C^*$ is isomorphic with the additive group of the n-tuples of integers, the homomorphism corresponding to the n-tuple $(p_1, \cdots, p_n)$ of integers being the homomorphism that sends $(c_1, \cdots, c_n)$ onto $c_1^{p_1} \cdots c_n^{p_n}$. Hence we see that the map $Q \cap S \to \mathrm{Hom}((A_S)^{B'}, C^*)$ sending each q onto η_q, where $\eta_q(\alpha) = \alpha^*(q)$, is surjective; if $\zeta \in \mathrm{Hom}((A_S)^{B'}, C^*)$ then ζ defines an element ζ^* of $\mathrm{Hom}((C^*)^n, C^*)$ such that $\zeta^*(\alpha^*(q_1), \cdots, \alpha^*(q_n)) = \zeta(\alpha)$, and if ζ^* corresponds as above to the n-tuple $(p_1, \cdots, p_n)$ then ζ is the image of the element $q_1^{p_1} \cdots q_n^{p_n}$ of $Q \cap S$ under the above map.

Now let ϕ be any element of $\mathrm{Hom}(X_S, C^*)$. Then $\phi \circ \rho^{-1}$ belongs to $\mathrm{Hom}((A_S)^{B'}, C^*)$ and, by what we have just seen, there is an element q_ϕ in $Q \cap S$ such that $(\phi \circ \rho^{-1})(\alpha) = \alpha(q_\phi)q_\phi^{-1}$, for every element α of $(A_S)^{B'}$. Hence, for every element β of X_S, $\phi(\beta) = \rho(\beta)(q_\phi)q_\phi^{-1}$. But $\rho(\beta) = \beta h_\beta$, with $h_\beta \in H_S$, and the elements of Q are H-fixed. Hence we have actually $\phi(\beta) = \beta(q_\phi)q_\phi^{-1}$.

Now let f be any element of $\boldsymbol{R}(G)$, and write $f = \sum_i b_i q_i$, with b_i in B and q_i in Q. Since B is an affine structure for G in $\boldsymbol{R}(G)$, we know from the last statement of Theorem 2.2 that B' is precisely the $(A_S)^{B'}$-fixed part of S. Since $(A_S)^{B'}$ lies in the center of M, it follows that B' is X_S-stable. Since X_S is abelian and reductive, the X_S-module, U_i say, that is generated by b_i' is a direct sum of 1-dimensional X_S-modules Cu_{ij}. For β in X_S, let us write $\beta(u_{ij}) = \phi_{ij}(\beta)u_{ij}$. Then $\phi_{ij} \in \mathrm{Hom}(X_S, C^*)$, and we know from the above that there is an element q_{ij} in $Q \cap S$ such that $\phi_{ij}(\beta) = \beta(q_{ij})q_{ij}^{-1}$, for every element β of X_S. We have $b_i' = \sum_j c_{ij}u_{ij}$, with c_{ij} in C. Now we have

$$f' = \sum_i b_i' q_i^{-1} = \sum_{i,j} c_{ij}u_{ij}q_i^{-1}$$
$$= \sum_{i,j} c_{ij}(u_{ij}q_{ij}^{-1})\,q_{ij}q_i^{-1}.$$

But each $u_{ij}q_{ij}^{-1}$ lies in $R(G)^X$, by the definition of the q_{ij}'s. Thus we have $f' \in \boldsymbol{R}(G)^X[Q]$, and we conclude that $\boldsymbol{R}(G) = \boldsymbol{R}(G)^X[Q]$.

In order to conclude that $\boldsymbol{R}(G)^X$ is a basic subalgebra of $\boldsymbol{R}(G)$, it remains only to show that Q is free over $\boldsymbol{R}(G)^X$. Since we may choose S

so as to contain any given finite subset of Q, it suffices to show that $S \cap Q$ is free over $\boldsymbol{R}(G)^X$.

Suppose this is not the case, and let $\sum_{i=1}^{n} f_i q_i = 0$ be a minimal non-trivial relation, with f_i in $\boldsymbol{R}(G)^X$ and q_i in $S \cap Q$. Operate with an element β of X to obtain $\sum_i f_i \beta(q_i)(1) q_i = 0$. On the other hand, multiply the original relation by $\beta(q_1)(1)$ and subtract the result from the last relation. This yields $\sum_i f_i(\beta(q_i)(1) - \beta(q_1)(1)) q_i = 0$ By the minimality of the original relation, we must therefore have $\beta(q_i)(1) = \beta(q_1)(1)$, for all β in X_S and all i. We have seen above that $\beta(q) = \rho(\beta)(q)$, for all elements β of X_S and all elements q of $Q \cap S$. Hence the last result gives $\alpha(q_i)(1) = \alpha(q_1)(1)$, for all elements α of $(A_S)^{B'}$. In the notation used above, this means that $\alpha^*(q_i) = \alpha^*(q_1)$, for all elements α of $(A_S)^{B'}$. Since the map $\alpha \to \alpha^*$ sends $(A_S)^{B'}$ onto all of $\mathrm{Hom}(Q \cap S, C^*)$, we may therefore conclude that $q_i = q_1$. Since our relation was minimal, this means that $n = 1$, which is impossible for a non-trivial relation. Hence $S \cap Q$ is free over $\boldsymbol{R}(G)^X$, and we have shown that $\boldsymbol{R}(G)^X$ is a basic subalgebra of $\boldsymbol{R}(G)$.

This implies that $A^{\boldsymbol{R}(G)^X} \cap G = (1)$. Since $A = X \cdot G$ and $X \subset A^{\boldsymbol{R}(G)^X}$, we conclude that $X = A^{\boldsymbol{R}(G)^X}$, and Theorem 4.1 is proved.

THEOREM 4.2. *The affine structures for G in $\boldsymbol{R}(G)$ are precisely the left stable basic subalgebras of $\boldsymbol{R}(G)$.*

Proof. We know already from Proposition 3.1 that every left stable basic subalgebra is an affine structure. Now suppose that P is an affine structure for G in $\boldsymbol{R}(G)$. Then we have from Theorem 2.2 that A is the semidirect product $A^{P'} \cdot G$. Now $A^{P'}$ is evidently a pro-algebraic subgroup of A. Also, since the commutator subgroup of A lies in G, $A^{P'}$ is abelian. In proving Theorem 3.4, we saw that $(A^{P'})_S$ is reductive, for every fully stable finitely generated subalgebra S containing P'. Hence it is clear that $A^{P'}$ satisfies the conditions imposed on X in Theorem 4.1, so that $\boldsymbol{R}(G)^{A^{P'}}$ is a basic subalgebra of $\boldsymbol{R}(G)$. By Theorem 2.2, $\boldsymbol{R}(G)^{A^{P'}} = P'$. Thus P' is a basic subalgebra of $\boldsymbol{R}(G)$, whence also P is a basic subalgebra of $\boldsymbol{R}(G)$. This completes the proof of Theorem 4.2.

In order to proceed, we need some further information on reductive groups. The abelian reductive complex analytic groups are the direct products of copies of C^*, and we shall call them *complex toroids.*

The first lemma below is simply the known complexified version of the standard theorem on maximal toroidal subgroups of compact real analytic groups.

LEMMA 4.3. *Let H be a reductive complex analytic group, and let T be a maximal complex toroid in H. Then T coincides with its centralizer in H.*

Proof. Let T^* denote the centralizer of T in H, and let X be the maximum compact subgroup of T. Choose a maximal compact subgroup Y of H containing X and a maximal compact subgroup of T^*. Then Y is a compact real analytic group, and X is a toroidal subgroup of Y. Since T is a maximal complex toroid in H, it is clear that X must be a *maximal* toroidal subgroup of Y. By the standard theorem on maximal toroidal subgroups of compact real analytic groups, X therefore coincides with its centralizer in Y. Hence X is a maximal compact subgroup of T^*.

Hence the Lie algebra $X^{\cdot}$ of X is a Cartan subalgebra of the Lie algebra $Y^{\cdot}$ of Y, and therefore its complexification is a Cartan subalgebra of the complexification of $Y^{\cdot}$, i. e., the Lie algebra $T^{\cdot}$ of T is a Cartan subalgebra of the Lie algebra $H^{\cdot}$ of H.

Hence we have a decomposition $H^{\cdot} = T^{\cdot} + \sum_{\alpha \neq 0} S_\alpha$, where the S_α's are the 1-dimensional root spaces for the non-zero roots α of $T^{\cdot}$. Clearly, each S_α is stable under the image of T^* under the adjoint representation of H. Thus, if Z denotes the center of H, T^*/Z is a fully reducible algebraic linear group, and hence is reductive. On the other hand, the center of a reductive complex analytic group is reductive [7, p. 97], so that Z is reductive. Now the same argument we have already used in proving Theorem 3.4 shows that T^* has a full compact subgroup. Since X is a maximal compact subgroup of T^*, we conclude therefore that X is a full compact subgroup of T^*, whence we must have $T^* = T$, as was to be proved.

LEMMA 4.4. *Let M be a reductive complex analytic group, H a reductive complex analytic subgroup of M containing the commutator subgroup of M. Then there is a complex toroid X in M such that M is the semidirect product $X \cdot H$.*

Proof. Let M' denote the commutator subgroup of M, and let $Z(M)$ denote the center of M. Then, since the Lie algebra of M is reductive, we have $M = M'Z(M)$. Choose a maximal complex toroid T_H in H and a maximal complex toroid T_M in M that contains T_H. Then we have a direct decomposition $T_M = T_H \times X$, where X is a complex toroid. By Lemma 4.3, applied to the maximal complex toroid T_M of M, we have $Z(M) \subset T_M$. Since H contains M', we have therefore $XH = M$. Now $X \cap H \subset T_M \cap H$, and $T_M \cap H$ is an abelian subgroup of H containing T_H. By Lemma 4.3, we have therefore $T_M \cap H = T_H$. Hence $X \cap H = X \cap T_M \cap H = X \cap T_H = (1)$. This establishes Lemma 4.4.

THEOREM 4.5. *Every fully stable finitely generated subalgebra of* $\mathbf{R}(G)$ *that separates the points of* G *and contains* $\mathrm{Hom}(G, C)$ *also contains a left stable basic subalgebra of* $\mathbf{R}(G)$.

Proof. Let S be a subalgebra of $\mathbf{R}(G)$ satisfying the conditions of the theorem. Write G as a semidirect product $H \cdot K$, where K is a nucleus (in the sense of [8]) and H a maximal reductive subgroup of G. Let M be a maximal fully reducible subgroup of A_S such that $H_S \subset M$. Since K_S is normal in G_S, it is also normal in the algebraic group hull A_S of G_S. Hence $M \cap K_S$ is normal in the fully reducible group M of automorphisms of S. It follows that S is semisimple as an $(M \cap K_S)$-module, which implies that $M \cap K_S$ must leave the elements of $\mathrm{Hom}(G, C)$ fixed. Now it is known from [8] that the kernel of the representation of K_S on $C + \mathrm{Hom}(G, C)$ is N_S, where N denotes the radical of the commutator subgroup of G. Thus $M \cap K_S \subset N_S$, which implies that the representation of $M \cap K_S$ on every finite-dimensional stable subspace of S is unipotent. Since these representations are also semisimple, we conclude that $M \cap K_S = (1)$. Hence we have $M \cap G_S = H_S \cdot (M \cap K_S) = H_S$. Since the commutator subgroup of A_S lies in G_S, this shows that H_S contains the commutator subgroup of M. Hence we may apply Lemma 4.4 to conclude that there is a complex toroid X in M such that M is the semidirect product $X \cdot H_S$.

By Lemma 3.3, we have $A_S = MG_S$, so that $A_S = (X \cdot H_S) G_S = XG_S$. Moreover, $X \cap G_S = X \cap M \cap G_S = X \cap H_S = (1)$. Thus A_S is the semidirect product $X \cdot G_S$. Since X is a fully reducible algebraic subgroup of A_S, the algebraic variety $X \backslash A_S$ is affine, and the algebra of all polynomial functions on $X \backslash A_S$ is $(S^X)'$. Since the canonical map $G \to X \backslash A_S$ is an isomorphism of complex analytic manifolds, $(S^X)'$ is therefore an affine structure for G in $\mathbf{R}(G)$. By Theorem 4.2, it is therefore a left stable basic subalgebra of $\mathbf{R}(G)$, so that Theorem 4.5 is proved.

Note. The algebraic hulls A_S *of* G *with* S *as in Theorem* 4.5 *are precisely all those algebraic hulls in which* G *is a semidirect factor.* Indeed, if G^* is any algebraic hull of G such that G^* is a semidirect product $X \cdot G$ then it follows from Lemma 3.3 that X has a full compact subgroup, and hence that X is a complex toroid. The last part of the proof of Theorem 4.5 now shows that G^* may be identified with an A_S, where S satisfies the conditions of Theorem 4.5. Moreover, we see from Theorem 4.5 and its proof that *these algebraic hulls* G^* *are also characterized by the property that the elements of* $\mathrm{Hom}(G, C)$ *are the composites of the injection* $G \to G^*$ *with the rational homomorphisms of* G^* *into* C.

5. **Lifting of quasi-affine structures.** We shall say that a quasi-affine

structure on $L\backslash G$ is *liftable* to G if there is an affine structure on G such that the canonical map $G \to L\backslash G$ is a morphism of algebraic varieties.

THEOREM 5.1. *Let L be a closed complex Lie subgroup of G, and suppose that $L\backslash G$ has a quasi-affine structure P in $\boldsymbol{R}(G)$. Then this structure is liftable to G if and only if $A^{P'}$ is a semidirect product $X \cdot L$, where X is an abelian pro-algebraic subgroup of A all whose elements are semisimple automorphisms of $\boldsymbol{R}(G)$.*

Proof. Suppose first that P is liftable to G. Then we know from Theorem 4.2 that there is a left stable basic subalgebra B of $\boldsymbol{R}(G)$ that is an affine structure for G in $\boldsymbol{R}(G)$ such that $P \subset B$. Moreover, we know from the proof of Theorem 4.2 that A is the semidirect product $A^{B'} \cdot G$, and and that $A^{B'}$ is an abelian pro-algebraic subgroup of A consisting entirely of semisimple elements. Since $P \subset B$, we have $P' \subset B'$ and hence $A^{B'} \subset A^{P'}$. Hence $A^{P'} = A^{B'} \cdot (A^{P'} \cap G) = A^{B'} \cdot L$. Thus the condition of our theorem is satisfied, with $X = A^{B'}$.

Now suppose that $A^{P'} = X \cdot L$, as in the statement of the theorem. Since $A^{P'}G = A$ and $A^{P'} \cap G = L$, we have that A is the semidirect product $X \cdot G$. Put $B = (\boldsymbol{R}(G)^X)'$. By Theorem 4.1, B is a left stable basic subalgebra of $\boldsymbol{R}(G)$, and $X = A^{B'}$. Since $X \subset A^{P'}$, we have $\boldsymbol{R}(G)^{A^{P'}} \subset B'$, i.e., $P' \subset B'$, whence $P \subset B$. This means that if G is equipped with the affine structure defined by B then the canonical map $G \to L\backslash G$ is a morphism of algebraic varieties. Our proof of Theorem 5.1 is now complete.

Suppose that P is a quasi-affine structure for $L\backslash G$ in $\boldsymbol{R}(G)$ that is liftable to G. Denote the pro-algebraic hull of L in A by L^*. Evidently, $L^* \subset A^{P'}$. Let $A^{P'} = X \cdot L$ be a semidirect product decomposition as is given by Theorem 5.1. This yields a semidirect product decomposition $L^* = (X \cap L^*) \cdot L$. Clearly, $X \cap L^*$ is still a pro-algebraic subgroup of A. From the last part of the proof of Theorem 5.1, we know that $X = A^{B'}$, where B is a left stable basic subalgebra of $\boldsymbol{R}(G)$. Let S be a fully stable finitely generated subalgebra of $\boldsymbol{R}(G)$ containing B'. Then we have $X_S \cap G_S = (1)$. Hence $(L^*)_S$ is the semidirect product $(X \cap L^*)_S \cdot L_S$. Since $(L^*)_S$ and $(X \cap L^*)_S$ are algebraic groups, they have only a finite number of connected components. Hence also L_S has only a finite number of connected components, so that L/L_1 is finite. Thus if $L\backslash G$ has a quasi-affine structure that is liftable to G then L/L_1 is finite.

This enables us to give an example of a non-liftable quasi-affine structure. Let G be the group of the matrices $\begin{pmatrix} e^a & 0 \\ b & e^{ia} \end{pmatrix}$, where a and b range over C. Let G^* be the algebraic hull of G for the identity representation;

it consists of the matrices $\begin{pmatrix} u & 0 \\ b & v \end{pmatrix}$, where u and v range over C^*, and b over C. Let X be the algebraic subgroup of G^* that is defined by putting $v=1$ and $b=0$. Since X is fully reducible, the algebraic variety $X\backslash G^*$ is affine. Since $XG=G^*$, we have an induced affine structure on $(X \cap G)\backslash G$. Now $X \cap G$ is an infinite discrete group, consisting of the matrices $\begin{pmatrix} e^{2\pi n} & 0 \\ 0 & 1 \end{pmatrix}$, where n ranges over the rational integers. Hence this affine structure of $(X \cap G)\backslash G$ is not liftable to G.

We shall show that the condition that L/L_1 be finite is actually sufficient for every quasi-affine structure on $L\backslash G$ to be liftable to G. For doing this, we require the following lemma concerning reductive groups.

LEMMA 5.2. *Let H be a reductive complex Lie group, and let K be a closed normal complex Lie subgroup of H such that H/K has a faithful finite-dimensional complex analytic representation. Then K is reductive.*

Proof. We use the basic facts concerning reductive groups that we recalled in Section 3. We may identify H with a fully reducible algebraic linear group, and then every finite-dimensional complex analytic representation of H is a rational representation. Hence the assumption of our lemma implies that K is the kernel of a rational representation, and hence a normal algebraic subgroup of the fully reducible algebraic linear group H. Thus K is also a fully reducible algebraic linear group, which implies that K is reductive.

THEOREM 5.3. *Let L be a closed complex Lie subgroup of G. If $L\backslash G$ has a quasi-affine structure that is liftable to G then L/L_1 is finite. Conversely, if L/L_1 is finite then every quasi-affine structure on $L\backslash G$ is liftable to G.*

Proof. The first part of this theorem has already been proved, and we shall now prove the converse. Thus we assume that L/L_1 is finite, and that P is a quasi-affine structure for $L\backslash G$ in $\boldsymbol{R}(G)$. Choose a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ containing an affine carrier for $L\backslash G$. By Theorem 2.2, we have $(A^{P'})_S G_S = A_S$ and $(A^{P'})_S \cap G_S = L_S$. Also, the commutator subgroup of $(A^{P'})_S$ lies in G_S, and hence in L_S. Hence L_S is normal in $(A^{P'})_S$ and $(A^{P'})_S/L_S$ is canonically isomorphic with A_S/G_S. We know from the note at the end of Section 4 that if S is taken large enough (so as to contain a left stable basic subalgebra of $\boldsymbol{R}(G)$) then A_S/G_S is reductive. Thus, if S is so chosen, $(A^{P'})_S/L_S$ is reductive. Being an algebraic group, $(A^{P'})_S$ has only a finite number of connected components,

so that the theory of maximal compact subgroups applies to it. Since $L_S/(L_S)_1$ is finite, the canonical image of every maximal compact subgroup of $(A^{P'})_S$ contains a maximal, and hence full, compact subgroup of $(A^{P'})_S/L_S$. It follows that if M is a maximal fully reducible subgroup of $(A^{P'})_S$ then $ML_S = (A^{P'})_S$. Take M such that it contains a maximal reductive subgroup H_S of L_S. The kernel of the canonical epimorphism $M \to (A^{P'})_S/L_S$ is $M \cap L_S$. By Lemma 5.2, $M \cap L_S$ is therefore reductive. Since it contains H_S, we have therefore $M \cap L_S = H_S$.

We have $(H_S)_1 \subset M_1$, and the commutator subgroup of M_1 lies in $(H_S)_1$. Hence we may apply Lemma 4.4 to conclude that there is a complex toroid X in M_1 such that M_1 is the semidirect product $X \cdot (H_S)_1$. Now observe that, since $(A^{P'})_S G_S = A_S$, we have already $((A^{P'})_S)_1 G_S = A_S$. Also, $((A^{P'})_S)_1 = M_1(L_S)_1 = X(L_S)_1$. Hence $XG_S = A_S$. On the other hand, $X \cap G_S = X \cap H_S$, which is finite, because $X \cap (H_S)_1 = (1)$.

The desired affine structure on G that is to exhibit the liftability of the given quasi-affine structure on $L\backslash G$ would be induced from $X\backslash A_S$ if we hand $X \cap G_S = (1)$ rather than only $X \cap G_S$ finite. We shall remedy this defect in the above construction of X by passing to a suitable covering group of A_S.

Let us construct the appropriate semidirect product $X \cdot G_S$ as a complex analytic group such that the multiplication in A_S induces a complex analytic epimorphism $X \cdot G_S \to A_S$. Let U denote the maximum unipotent normal subgroup of A_S, and let J be a maximal fully reducible subgroup of A_S such that $X \subset J$. Denote by U^0 and J^0 the connected components of the identity in the complete inverse images of U and J in $X \cdot G_S$, respectively. Since $A_S = J \cdot U$, we have then $X \cdot G_S = J^0 U^0$. Now consider the restriction to U^0 of the canonical map $X \cdot G_S \to J\backslash A_S$. This exhibits U^0 as a covering space of the simply connected space $J\backslash A_S$. Hence the map $U^0 \to J\backslash A_S$ is a homeomorphism. Hence $J^0 \cap U^0 = (1)$, so that $X \cdot G_S$ is the semidirect product $J^0 \cdot U^0$, the map $X \cdot G_S \to A_S$ sends U^0 isomorphically onto U, and J^0 is the *complete* inverse image of J. Let E denote the kernel of the epimorphism $X \cdot G_S \to A_S$. By what we have just seen, $E \subset J^0$, and evidently E is finite. Since J^0/E is reductive (being isomorphic with J), so is therefore J^0.

Hence J^0 has a unique structure of an algebraic group such that every finite-dimensional complex analytic representation of J^0 is rational. On the other hand, U^0 may be equipped with the structure of a unipotent algebraic group, isomorphic as such with U. The map that attaches to each element of J^0 the conjugation it effects on U^0 is a rational homomorphism of J^0 into the group of all rational automorphisms of U^0, as is clear by comparing this action of J^0 on U^0 with the action of J on U within A_S. Hence the semi-

direct product $J^0 \cdot U^0$ may be equipped with the structure of an algebraic group with which it becomes the semidirect product, in the sense of algebraic groups, of the algebraic groups J^0 and U^0. The requisite technical details will be found in [4, Section 6]. We regard our semidirect product $X \cdot G_S$ as this algebraic group $J^0 \cdot U^0$, and it is clear from the construction that then our covering $X \cdot G_S \to A_S$ is a rational group epimorphism.

Now let $G_S{}^*$ denote the algebraic group hull of G_S in $X \cdot G_S$. Then the image of $G_S{}^*$ in A_S is an algebraic subgroup of A_S containing G_S, and hence coincides with A_S. Hence $G_S{}^*E = X \cdot G_S$, so that the index of $G_S{}^*$ in $X \cdot G_S$ is finite. Since $X \cdot G_S$ is connected, and hence irreducible as an algebraic group, it follows that $G_S{}^* = X \cdot G_S$, i. e., that G_S is algebraically dense in $X \cdot G_S$. Hence, if $\boldsymbol{P}(X \cdot G_S)$ denotes the algebra of all polynomial functions on the algebraic group $X \cdot G_S$, the restriction map $\boldsymbol{P}(X \cdot G_S) \to \boldsymbol{P}(X \cdot G_S)_{G_S}$ is an isomorphism. We may identify $\boldsymbol{P}(X \cdot G_S)_{G_S}$ with the corresponding subalgebra, T say, of $\boldsymbol{R}(G)$. Clearly, T is a finitely generated fully stable subalgebra of $\boldsymbol{R}(G)$. Moreover, since the epimorphism $X \cdot G_S \to A_S$ is rational, we have $S \subset T$.

Now $X \cdot G_S$ may be viewed as the group of all proper automorphisms of $\boldsymbol{P}(X \cdot G_S)$ and hence may be identified with the corresponding group of automorpisms of T. This is the group of all proper automorphisms of T, and thus coincides with A_T. Thus, if X_T is the isomorphic image of X in the group of the proper automorphisms of T, we have $A_T = X_T \cdot G_T$. The restriction epimorphism $A_T \to A_S$ sends X_T isomorphically onto X and G_T isomorphically onto G_S. Since S contains an affine carrier for $L\backslash G$, $(A^{P'})_T$ is the complete inverse image of $(A^{P'})_S$ in A_T. Hence we have $X_T \subset (A^{P'})_T$.

Since X_T is a fully reducible algebraic subgroup of A_T, the algebraic variety $X_T\backslash A_T$ is affine. As a complex analytic manifold, $X_T\backslash A_T$ is canonically isomorphic with G_T, and thus with G, because we have taken S so as to separate the points of G. Thus we have an induced affine structure on G such that the corresponding affine structure for G in $\boldsymbol{R}(G)$ is $(T^{X_T})'$. Since $X_T \subset (A^{P'})_T$, we have $T^{A^{P'}} \subset T^{X_T}$, i. e., $P' \cap T \subset T^{X_T}$, whence $P \cap T \subset (T^{X_T})'$. Now S, and hence T, contains an affine carrier for $L\backslash G$ in $\boldsymbol{R}(G)$. Hence the intersection with $\boldsymbol{R}(G)$ of the field of fractions of $P \cap T$ coincides with P. On the other hand, since $(T^{X_T})'$ is an affine structure for G in $\boldsymbol{R}(G)$, the intersection with $\boldsymbol{R}(G)$ of its field of fractions coincides with $(T^{X_T})'$. Hence we have $P \subset (T^{X_T})'$, so that the canonical map $G \to L\backslash G$ is a morphism of algebraic varieties. Thus we have proved that the given quasi-affine structure for $L\backslash G$ is liftable to G, so that Theorem 5.3 is established.

6. Separation. The object of this section is to prove the following result.

THEOREM 6.1. *Let L be a closed complex Lie subgroup of G such that $(R(G)^L)'$ separates the points of $L\backslash G$. Then there is actually a finite subset of $(R(G)^L)'$ that separates the points of $L\backslash G$.*

Proof. We show first that LG' is closed in G, where G' denotes the commutator subgroup of G. We recall that G' is a semidirect product of its radical and a reductive group, and that every finite-dimensional complex analytic representation of G is unipotent on the radical of G'. It follows that G' has the structure of an algebraic group such that the restriction to G' of every finite-dimensional complex analytic representation of G is a rational representation of G'. In particular, if V is any finite-dimensional left G-stable subspace of $\boldsymbol{R}(G)$ then G'_V is an algebraic subgroup of the full linear group on V, and the canonical epimorphism $G' \to G'_V$ is a rational representation of G'.

Let T be a finite subset of $\boldsymbol{R}(G)^L$, and let V be the smallest left G-stable subspace of $\boldsymbol{R}(G)$ containing T. Then $(G'_V)^T$ is an algebraic subgroup of G'_V, and its complete inverse image in G' is $(G')^T$. Hence $(G')^T$ is an algebraic subgroup of G'. The intersection of the family of groups G^T, as T ranges over all finite subsets of $\boldsymbol{R}(G)^L$, is evidently L. Hence the intersection of the family of groups $(G')^T$ is $L \cap G'$. Since these are algebraic subgroups of the algebraic group G', it follows that there is a T_0 among these T's such that $(G')^T = L \cap G'$ for all T's containing T_0.

Now consider the family of algebraic subgroups $(A_V)^T G'_V$ of A_V. Since the canonical map $G \to A_V$ is continuous and $(A_V)^T G'_V$ is closed in A_V, the inverse image of $(A_V)^T G'_V$ in G is closed in G. The kernel of the canonical homomorphism $G \to A_V$ is evidently contained in G^T. Hence the inverse image of $(A_V)^T G'_V$ in G is $G^T G'$. Thus $G^T G'$ is closed in G. We claim that the intersection of the family of groups $G^T G'$, as T ranges over the finite subsets of $\boldsymbol{R}(G)^L$, is LG'. In order to see this, suppose that $T_0 \subset T$, $x \in G^T$, $x_0 \in G^{T_0}$, $y \in G'$, $y_0 \in G'$, and $xy = x_0 y_0$. Then we have $x^{-1}x_0 = yy_0^{-1} \in (G')^{T_0} \subset L$, whence $x_0 \in G^T$. Hence, if $x_0 y_0$ belongs to the intersection of the family of groups $G^T G'$ then x_0 belongs to the intersection of the family of groups G^T, i. e., $x_0 \in L$, so that $x_0 y_0 \in LG'$. Thus our above claim is established, and we may conclude that LG' is closed in G.

Since $(A_V)^T G'_V$ is a normal algebraic subgroup of A_V, we know that if W is the smallest fully stable subalgebra of $\boldsymbol{R}(G)$ containing V (which is the algebra of all polynomial functions on A_V) then the $(A_V)^T G'_V$-fixed part of W separates the points of the factor group $((A_V)^T G'_V)\backslash A_V$. The kernel of the canonical epimorphism of G onto this last group is $G^T G'$. Hence we may conclude that the $(A_V)^T G'_V$-fixed part of W separates the

points of $(G^TG')\backslash G$. A fortiori, the G^TG'-fixed part of $\boldsymbol{R}(G)$ separates the points of $(G^TG')\backslash G$. Since the intersection of the family of groups G^TG' is LG', it follows that the LG'-fixed part of $\boldsymbol{R}(G)$ separates the points of $(LG')\backslash G$.

Now this last group is an abelian complex analytic group. By [7, Lemma 2.1], it is of the form $E\times U$, where U is a vector group and E has a full compact subgroup. This last fact implies, via the Peter-Weyl Theorem and [7, Lemma 2.2], that the points of E are already separated by a finite subset of $\boldsymbol{R}(G)^{LG'}$. On the other hand, the points of U are separated by a finite subset of $\mathrm{Hom}((LG')\backslash G, C)$, and the elements of $\mathrm{Hom}((LG')\backslash G, C)$ are constant on the cosets of E. Hence we see that the points of $(LG')\backslash G$ are already separated by some finite subset X of $\boldsymbol{R}(G)^{LG'}$.

Now $(\boldsymbol{R}(G)^L)'$ separates the points of $(L\cap G')\backslash G'$, and $L\cap G'$ is an algebraic subgroup of G' (as we have seen above). Hence there is a finite subset Y of $(\boldsymbol{R}(G)^L)'$ that already separates the points of $(L\cap G')\backslash G'$, and thus the points of $L\backslash(LG')$. Now it is easy to see that the finite subset $X\cup Y$ of $(\boldsymbol{R}(G)^L)'$ separates the points of $L\backslash G$, so that Theorem 6.1 is proved.

7. Existence of quasi-affine structures. Let α denote the adjoint representation of G on its Lie algebra $G^{\cdot}$. We observe that *the algebraic group hull $\alpha(G)^*$ of $\alpha(G)$ in the full linear group on $G^{\cdot}$ lies in the canonical image of the group $\boldsymbol{A}(G)$ of all complex analytic automorphisms of G.* In order to see this, consider any algebraic group hull G^* of G. Since G is normal in G^*, we have a complex analytic homomorphism $G^*\rightarrow\boldsymbol{A}(G)$ sending each element of G^* onto the conjugation it effects on G. The composite of this homomorphism with the canonical monomorphism of $\boldsymbol{A}(G)$ into the group of automorphisms of $G^{\cdot}$ is the $G^{\cdot}$-part of the adjoint representation of G^*. Hence it is clear that the image of G^* under this canonical composite homomorphism is precisely $\alpha(G)^*$, which proves our above assertion.

Let L be a subset of G. Then we shall denote by $\boldsymbol{S}(L)$ the subgroup of $\alpha(G)^*$ consisting of all those elements of $\alpha(G)^*$ which correspond to elements of $\boldsymbol{A}(G)$ leaving L stable. Clearly, if G^* is any algebraic group hull of G then $S(L)$ is precisely the canonical image in $\alpha(G)^*$ of the normalizer of L in G^*. The normalizer of L in G will be denoted by $\boldsymbol{N}(L)$. Now we are ready for the statement of the main existence theorem.

THEOREM 7.1. *Let L be a closed complex Lie subgroup of G. Then $L\backslash G$ has a quasi-affine structure if and only if the following three conditions are satisfied*: (1) $(\boldsymbol{R}(G)^L)'$ *separates the points of* $L\backslash G$; (2) $\boldsymbol{S}(L)\alpha(G)=\alpha(G)^*$; (3) $\boldsymbol{N}(L)/L$ *has only a finite number of connected components.*

The proof of Theorem 7.1 is rather long, and we shall divide it into several parts. First, we shall prove the necessity of the three conditions. Evidently, (1) is a necessary condition. In order to prove the necessity of conditions (2) and (3), suppose that $L\backslash G$ has a quasi-affine structure. Then we know from Theorem 2.2 that there is an algebraic hull G^* of G and an algebraic subgroup F of G^* such that $FG = G^*$, $F \cap G = L$, and the analytic manifold $L\backslash G$ is canonically isomorphic with $F\backslash G^*$. Let α^* denote the adjoint representation of G^* on the Lie algebra $G^\cdot$ of G. Then we have $\alpha^*(F)\alpha(G) = \alpha^*(G^*) = \alpha(G)^*$. Since the commutator subgroup of G^* lies in G, and since $F \cap G = L$, it is clear that F lies in the normalizer of L in G^*. Hence $\alpha^*(F) \subset \boldsymbol{S}(L)$, so that the above yields (2). Let $\boldsymbol{N}_G$ stand for normalizer in G and $\boldsymbol{N}_{G^*}$ for normalizer in G^*. We have just seen that $F \subset \boldsymbol{N}_{G^*}(L)$. Since $FG = G^*$, we have therefore $\boldsymbol{N}_{G^*}(L) = F\boldsymbol{N}_G(L)$. This means that, under the analytic manifold isomorphism between $L\backslash G$ and $F\backslash G^*$, the subspace $L\backslash \boldsymbol{N}_G(L)$ of $L\backslash G$ corresponds to the subspace $F\backslash \boldsymbol{N}_{G^*}(L)$ of $F\backslash G^*$. Thus $L\backslash \boldsymbol{N}_G(L)$ is homeomorphic with $F\backslash \boldsymbol{N}_{G^*}(L)$. Since $\boldsymbol{N}_{G^*}(L)$ is an algebraic subgroup of G^*, $F\backslash \boldsymbol{N}_{G^*}(L)$ has only a finite number of connected components. The same is therefore true for $L\backslash \boldsymbol{N}_G(L)$, and $L\backslash \boldsymbol{N}_G(L) = \boldsymbol{N}_G(L)/L$. Thus we have shown that condition (3) is necessary.

Next we show that, *in proving the sufficiency of the conditions, we may assume that L has only a finite number of connected components.* Suppose that L satisfies the conditions of Theorem 7.1. By condition (1) and Theorem 6.1, we can then find a finitely generated fully stable subalgebra S of $\boldsymbol{R}(G)$ such that $(S^L)'$ separates the points of $L\backslash G$ and S separates the points of G. Put $G^* = A_S$, and let L^* denote the algebraic group hull of L in G^*. Then we have $(L^*)_1 \subset (L^*)_1 L \subset L^*$, which shows that $(L^*)_1 L$ is an algebraic subgroup of L^*, whence $(L^*)_1 L = L^*$. Hence we have $L^* G = (L^*)_1 G$, so that $L^* G$ is a complex analytic subgroup of G^*. Since $(S^L)'$ separates the points of $L\backslash G$, the fixer of S^L in G coincides with L, whence we have $L^* \cap G = L$. Hence the injection $G \to L^* G$ induces a bijective analytic map of $L\backslash G$ onto $L^*\backslash (L^* G)$; actually, it follows from the standard facts concerning analytic homogeneous spaces that such a map is an isomorphism of analytic manifolds, but we do not need this result. Therefore, it suffices, for the reduction, to prove that the pair $(L^* G, L^*)$ satisfies the conditions of Theorem 7.1, because L^* has only a finite number of connected components, and every quasi-affine structure on $L^*\backslash (L^* G)$ yields a quasi-affine structure on $L\backslash G$ via the above analytic bijection.

We have $S^L = S^{L^*}$. The elements of S^{L^*} may be regarded as functions on $L^* G$, and, as such, they evidently belong to $\boldsymbol{R}(L^* G)^{L^*}$. We claim that $(S^{L^*})'$ separates the points of $L^*\backslash (L^* G)$, or, equivalently, that the fixer in

L^*G of S^{L^*} coincides with L^*. Let x be an element of this fixer. Since $L^*G = GL^*$, we may write $x = uv$, with u in G and v in L^*. Then u evidently belongs to the fixer of S^L in G, so that $u \in L$ and $x \in L^*$, which proves our claim. Thus condition (1) of Theorem 7.1 holds for the pair (L^*G, L^*).

From condition (2) for (G,L), we have $\mathbf{N}_{G^*}(L)\,G = G^*$. Evidently, $\mathbf{N}_{G^*}(L) = \mathbf{N}_{G^*}(L^*)$. Hence we have $\mathbf{N}_{G^*}(L^*)L^*G = \mathbf{N}_{G^*}(L)\,G = G^*$. Considering the adjoint representation of L^*G, we see from this that condition (2) holds for the pair (L^*G, L^*).

Now observe that

$$\begin{aligned}\mathbf{N}_{L^*G}(L^*) &= \mathbf{N}_{G^*}(L^*) \cap (L^*G)\\ &= \mathbf{N}_{G^*}(L) \cap (L^*G) = L^*(\mathbf{N}_{G^*}(L) \cap G) = L^*\mathbf{N}_G(L).\end{aligned}$$

Hence $\mathbf{N}_{L^*G}(L^*)/L^*$ is isomorphic with $\mathbf{N}_G(L)/L$. Hence condition (3) for (G,L) implies condition (3) for (L^*G, L^*). This completes the reduction to the case where L has finite component group.

The proof of the sufficiency of the conditions of Theorem 7.1 will appeal to the following simple fact, which we exhibit as a proposition for later reference.

PROPOSITION 7.2. *Let X and Y be closed complex Lie subgroups of G such that the set XY coincides with G and both $X\backslash G$ and $Y\backslash G$ have quasi-affine structures. Then also $(X \cap Y)\backslash G$ has a quasi-affine structure.*

Proof. Let a and b be arbitrary elements of G. Since $XY = G$, we have $ab^{-1} = xy$, with x in X and y in Y. Now $x^{-1}a = yb$ and thus is a common representative in G for the cosets Xa and Yb Thus the canonical map $G \to (X\backslash G) \times (Y\backslash G)$ is surjective, and hence induces a bijective analytic map $(X \cap Y)\backslash G \to (X\backslash G) \times (Y\backslash G)$. Since a direct product of two quasi-affiine algebraic varieties is still quasi-affine, the given quasi-affine structures on $X\backslash G$ and $Y\backslash G$ yield a quasi-affine structure on $(X\backslash G) \times (Y\backslash G)$ and hence, by the above analytic bijection, a quasi-affine structure on $(X \cap Y)\backslash G$. This completes the proof of Proposition 7.2.

Now we return to the proof of Theorem 7.1. Assume that L satisfies the three conditions of this theorem and, moreover, that L has only a finite number of connected components. This last fact, together with condition (3), implies that $\mathbf{N}_G(L)$ has only a finite number of connected components.

By condition (1) and Theorem 6.1, we can find a finitely generated fully stable subalgebra S of $\mathbf{R}(G)$ such that the fixer in G of S^L coincides with L, S separates the points of G, and $\mathrm{Hom}(G, C) \subset S$. From condition (2), we have $\mathbf{N}_{A_S}(L_S)\,G_S = A_S$, whence we see that $\mathbf{N}_{A_S}(L_S)\,/\mathbf{N}_{G_S}(L_S)$ is isomorphic with A_S/G_S. Since $\mathbf{N}_{G_S}(L_S)$ has only a finite number of connected

components we may therefore apply the same argument we used in the proof of Theorem 5.3 to conclude that if E is any maximal connected fully reducible subgroup of the algebraic group $N_{A_S}(L_S)$ then $EN_{G_S}(L_S)=N_{A_S}(L_S)$, so that $EG_S=A_S$.

Now let M be a maximal fully reducible subgroup of A_S containing E. Since S contain $\mathrm{Hom}(G,C)$, we know from the proof of Theorem 4.5 that $M\cap G_S$ is a maximal reductive subgroup H_S of G_S. Since $EG_S=A_S$, we have $M=EH_S$. Now $A_S=M\cdot U$, where U is the maximum unipotent normal subgroup of A_S. Thus $A_S=EH_SU$. It follows from the conjugacy theorem for the maximal reductive subgroups of G that H_SU is normal in A_S. We have $E=E'Z(E)$, where E' is the commutator subgroup of E and $Z(E)$ is the connected component of the identity in the center of E. Since $E'\subset H_S$, we have therefore $A_S=Z(E)H_SU$.

Now let $L_S{}^*$ denote the algebraic hull of L_S in A_S. Then $L_S{}^*H_SU$ is an algebraic subgroup of A_S, so that its intersection with $Z(E)$ is an algebraic subgroup of $Z(E)$. Since $Z(E)$ is a complex toroid, the connected component of the identity in this intersection is a direct factor of $Z(E)$. Thus there is a complex toroid E^0 in $Z(E)$ such that $E^0L_S{}^*H_SU=A_S$ and $E^0\cap(L_S{}^*H_SU)$ is finite. We shall show that, by suitably enlarging S, we may arrange to have $E^0\cap(L_S{}^*H_SU)=(1)$.

Since $(L_1)_S{}^*$ is a connected normal algebraic subgroup of $N_{A_S}(L_S)$, we see that $E\cap(L_1)_S{}^*$ is a maximal fully reducible subgroup of $(L_1)_S{}^*$, so that we have $(L_1)_S{}^*=(E\cap(L_1)_S{}^*)\cdot Y$, where Y is the maximum unipotent normal subgroup of $(L_1)_S{}^*$. Now consider the canonical epimorphism $A_S\to A_S/(H_SU)$. The image is isomorphic with M/H_S, which we know from Lemma 4.4 to be a complex toroid. It follows that the canonical image of Y in $A_S/(H_SU)$ must be trivial, i.e., that $Y\subset H_SU$. Hence we have $(L_1)_S{}^*H_SU=((L_1)_S{}^*\cap E)H_SU$. We wish to show that this group actually coincides with $L_S{}^*H_SU$.

Note first that, since L/L_1 is finite, we have $L_S{}^*=L_S(L_1)_S{}^*$. Also, if N denotes the radical of the commutator subgroup of G then HN is a closed normal analytic subgroup of G, and G/HN is a vector group. Clearly, $(L_1HN)/HN$ is a complex analytic subgroup of the vector group G/HN. Hence it is closed in G/HN, and the factor group $G/(L_1HN)$ is also a vector group. Thus $(LHN)/(L_1HN)$ is a subgroup of a vector group. Since it is finite, it must therefore be trivial, so that $LHN=L_1HN$, Hence we have also $L_SH_SN_S=(L_1)_SH_SN_S$. Now $L_S{}^*H_SN_S$ is evidently the algebraic group hull of $L_SH_SN_S$ in A_S, and $(L_1)_S{}^*H_SN_S$ is the algebraic group hull of $(L_1)_SH_SN_S$ in A_S. Hence these two groups coincide. Since $N_S\subset U$, it follows that we have indeed $L_S{}^*H_SU=(L_1)_S{}^*H_SU=((L_1)_S{}^*\cap E)H_SU$.

Moreover, $((L_1)_S^* \cap E)H_S$ is a fully reducible algebraic subgroup, so that $L_S^*H_SU$ is the semidirect product $[((L_1)_S^* \cap E)H_S] \cdot U$. Now observe that E^0 normalizes each of these two factors of $L_S^*H_SU$. We may therefore construct the semidirect product $E^0 \cdot (L_S^*H_SU)$ and endow it with the unique structure of an algebraic group for which $E^0 \cdot [((L_1)_S^* \cap E)H_S]$ is a maximal fully reducible subgroup and U is the maximum unipotent normal subgroup. The multiplication in A_S defines a rational epimorphism of our semidirect product onto A_S. The kernel of this epimorphism is finite, because $E^0 \cap (L_S^*H_SU)$ is finite. Moreover, since this last intersection lies in $((L_1)_S^* \cap E)H_S$, the kernel of our epimorphism lies in $E^0 \cdot [((L_1)_S^* \cap E)H_S]$.

Let $G_S{}^0$ denote the connected component of the identity in the inverse image of G_S in $E^0 \cdot (L_S^*H_SU)$. Let K be a nucleus of G, in the sense of [8], and let $K_S{}^0$ denote the connected component of the identity in the inverse image of K_S. Since $G = H \cdot K$, we have $G_S = H_S \cdot K_S$ and hence $G_S{}^0 = H_SK_S{}^0$. Since K_S is simply connected, the covering $K_S{}^0 \to K_S$ is an isomorphism. We shall deduce from this that the covering $G_S{}^0 \to G_S$ is also an isomorphism. Let a be an element of the kernel of this last covering, and write $a = hk$, with h in H_S and k in $K_S{}^0$. If k_1 is the image of k in K_S, we have $hk_1 = 1$. But this implies that both h and k_1 are equal to 1. Since the map $K_S{}^0 \to K_S$ is injective, we conclude that $k = 1$, and so that $a = 1$. Thus the covering $G_S{}^0 \to G_S$ is an isomorphism.

Since G_S is algebraically dense in A_S, the algebraic group hull of $G_S{}^0$ in $E^0 \cdot (L_S^*H_SU)$ covers A_S. Hence it is of finite index in our semidirect product. Since this last is connected, we conclude therefore that $G_S{}^0$ is algebraically dense in $E^0 \cdot (L_S^*H_SU)$.

Now we proceed exactly as in the proof of Theorem 5.3 and replace S with a larger finitely generated fully stable subalgebra T of $\boldsymbol{R}(G)$, consisting of the restrictions to $G_S{}^0$ (which may be identified with G) of the polynomial functions on our semidirect product. Our semidirect product is now identified with A_T. Since L_S^* is an algebraic subgroup of our semidirect product, it thus becomes identified with the algebraic group hull L_T^* of L_T in A_T. The groups E^0 and U are now regarded as subgroups of A_T. Thus A_T is the semidirect product $E^0 \cdot (L_T^*H_TU)$.

Write D for the algebraic subgroup $E^0 \cdot L_T^*$ of A_T. We recall that H_TN_T is a normal algebraic subgroup of A_T; in fact, it is the kernel of the representation of A_T on $C + \mathrm{Hom}(G, C)$; [8, p. 125]. Hence $D_TH_TN_T$ is an algebraic subgroup of A_T. We wish to show that its intersection with G_T is connected.

In order to simplify the notation, we shall now identify G with G_T and accordingly write H for H_T and N for N_T. Let P be a maximal fully reducible subgroup of DHN that contains H. Since $A_T = (DHN)U$, it is

clear that P is also a maximal fully reducible subgroup of A_T. As before for M and G_S, we have $P \cap G = H$. By Lemma 3.3, $PG = A_T$. Now we have

$$H\backslash(G \cap (DHN)) = H\backslash(G \cap (PDN)) = (G \cap P)\backslash(G \cap (PDN)).$$

Since $A_S = PG$, we have $D \subset P(G \cap (PD))$. Hence $P(G \cap (PDN)) = PDN$. All the groups involved here are locally compact and separable. Hence the homogeneous space $(G \cap P)\backslash(G \cap (PDN))$ is homeomorphic with the homogeneous space $P\backslash(PDN)$. Since P is a maximal fully reducible subgroup of PDN, we see from the usual decomposition of the algebraic group PDN that $P\backslash(PDN)$ is connected. Thus $H\backslash(G \cap (DHN))$ is connected. Since H is connected, this implies that $G \cap (DHN)$ is connected, as we wished to show.

We have already seen above that LHN is a closed normal subgroup of G, and that $G/(LHN)$ is a vector group. Now LHN is evidently contained in $G \cap (DHN)$ and, since the last group is connected, $(G \cap (DHN))/(LHN)$ is a complex analytic subgroup of the vector group $G/(LHN)$. Hence there is a closed normal complex Lie subgroup V of G containing LHN and such that $G/(LHN)$ is the direct product of $V/(LHN)$ and $(G \cap (DHN))/(LHN)$.

Evidently, $G \cap (DHN) = (G \cap D)HN$. Hence we have

$$(G \cap D)V = (G \cap D)HNV = (G \cap (DHN))V = G.$$

On the other hand, $(G \cap D) \cap V \subset (G \cap (DHN)) \cap V = LHN$, so that $(G \cap D) \cap V \subset D \cap (LHN)$. But $D \cap (LHN) = L_T{}^* \cap (LHN) = L$. Thus $(G \cap D) \cap V = L$.

In view of Propostion 7.2, it will now suffice to show that both $V\backslash G$ and $(G \cap D)\backslash G$ have quasi-affine structures. Now $V\backslash G$ is canonically isomorphic with the direct vector group factor $(LHN)\backslash(G \cap (DHN))$ of the vector group $(LHN)\backslash G$, and hence has evidently an affine structure. Thus it remains only to show that $(G \cap D)\backslash G$ has a quasi-affine structure. Since $DG = PG = A_T$, the injection $G \to A_T$ induces an analytic bijection of $(G \cap D)\backslash G$ onto $D\backslash A_T$, so that it suffices to prove that the algebraic variety $D\backslash A_T$ is quasi-affine.

This amounts to showing that D is *observable* in A_T, in the sense of [1]. Since DHN is a normal algebraic subgroup of A_T, we know from [1, Th. 10] that DHN is observable in A_T. In view of the evident transitivity property of the notion of observability, it suffices therefore to show that D is observable in DHN, or, equivalently, that the algebraic variety $D\backslash(DHN)$, which we may identify with $(D \cap (HN))\backslash(HN)$, is quasi-affine. Since $L_T{}^* \cap G_T = L_T$, we have $D \cap (HN) = L \cap (HN)$. Now the restrictions to HN of the elements of $(T^L)'$ are polynomial functions on HN that are constant on the cosets of $L \cap (HN)$ and, since $(T^L)'$ separates the points of $L\backslash G$, they

separate the points of $(L \cap (HN))\backslash(HN)$. By [1, Th. 4], this implies that the algebraic variety $(L \cap (HN))\backslash(HN)$ is quasi-affine, so that our proof of Theorem 7.1 is complete.

8. Examples. In all the examples we shall exhibit below, condition (1) of Theorem 7.1 will be satisfied, in virtue of the following lemma.

LEMMA 8.1. *Let V be a complex vector group, T an arbitrary complex Lie group, $G = V \cdot T$ a semidirect product, with V normal in G. Let W be a complex vector subgroup of V. Then $(\boldsymbol{R}(G)^W)'$ separates the point of $W\backslash G$.*

Proof. Let f be any element of $\mathrm{Hom}(V, C)$. Define the function f^+ on G by setting $f^+(vt) = f(v)$, for all v in V and all t in T. Then, by [6, Prop. 2.4], f^+ belongs to $\boldsymbol{R}(G)$. Now suppose that $f(W) = (0)$. Then it is clear that f^+ lies in $(\boldsymbol{R}(G)^W)'$. As f ranges over all these elements of $\mathrm{Hom}(V, C)$, f^+ ranges over a subspace S of $(\boldsymbol{R}(G)^W)'$. Now suppose that v_1t_1 and v_2t_2 are elements of G such that $g(v_1t_1) = g(v_2t_2)$, for all elements g of S. Then $f(v_1) = f(v_2)$, for all elements f in $\mathrm{Hom}(G, C)$ such that $f(W) = (0)$. Hence $v_2 \in Wv_1$. Moreover, $(\boldsymbol{R}(G)^V)' \subset (\boldsymbol{R}(G)^W)'$, so that, if $g(v_1t_1) = g(v_2t_2)$ for all elements g of $\boldsymbol{R}(G)^W)'$ then $t_1 = t_2$. Hence we have then $Wv_2t_2 = Wv_1t_2 = Wv_1t_1$. Thus $(\boldsymbol{R}(G))'$ separates the points of $W\backslash G$, and Lemma 8.1 is proved.

Let V be a vector group, with basis (v_1, v_2, v_3). Let T be the additive group of the complex numbers, and define a representation of T on V such that, for every t in T, $t \cdot v_1 = e^{-t}v_1$, $t \cdot v_2 = e^t(v_2 + tv_3)$, $t \cdot v = e^t v_3$. Let G be the corresponding semidirect product $V \cdot T$. Let W be the 1-dimensional vector subgroup of V that is spanned by $v_1 + v_2$. We shall show that W *satisfies condition* (3) *of Theorem* 7.1, *but not* (2).

Evidently, $N_G(W) = V$, so that (3) is satisfied. Let α denote the adjoint representation of G. Since $\alpha(V)$ is unipotent, we have $\alpha(G)^* = \alpha(V)\alpha(T)^*$. We may identify V with its Lie algebra, and we consider, in this sense, the restriction $\alpha(G)^*{}_V$ of $\alpha(G)^*$ on V. This evidently coincides with $\alpha(T)^*{}_V$, which consists of the linear automorphism $\tau_{(a,b)}$, where $a \in C^*$ (the multiplicative group of the non-zero complex numbers), $b \in C$, and the images of v_1, v_2, v_3 under $\tau_{(a,b)}$ are, respectively, $a^{-1}v_1$, $av_2 + bv_3$, av_3. Hence $S(W)_V$ consists of the two automorphisms $\tau_{(1,0)}$ and $\tau_{(-1,0)}$. It is clear from this that $S(W)\alpha(G) \neq \alpha(G)^*$, i.e., that (2) is violated.

Now let W^+ denote the 2-dimensional vector subgroup of V that is spanned by $v_1 + v_2$ and v_3. We shall show that W^+ *satisfies* (2), *but not* (3).

Let F denote the stabilizer of W^+ in $\alpha(T)^*$. Then we have $S(W^+) = \alpha(V)F$. One sees immediately that F consists of those elements of $\alpha(T)^*$ whose restrictions to V are the automorphisms $\tau_{(1,b)}$ and $\tau_{(-1,b)}$, with arbitrary

b in C. Hence it is clear that $F\alpha(T))_V = \alpha(T)^*_V$. The restriction map $\alpha(T)^* \to \alpha(T)^*_V$ is evidently an isomorphism. Hence we have $F\alpha(T) = \alpha(T)^*$. Together with $S(W^+) = \alpha(V)F$, this gives $S(W^+)\alpha(G) = \alpha(G)^*$, i.e., (2) is satisfied. On the other hand, we have $N_G(W^+) = V \cdot T^0$, where T^0 consists of the elements t in T such that $e^{-t} = e^t$, i.e., of the integral multiples of πi. Hence the group of components of $N_G(W^+)/W^+$ is isomorphic with the infinite group T^0, so that (3) is violated.

Note that *both $W\backslash V$ and $V\backslash G$ have the structure of an algebraic linear group, and yet $W\backslash G$ does not have a quasi-affine structure.* This is in contrast with the situation for algebraic homogeneous spaces [1, end of Section 5].

Let H be the semidirect product $G \cdot C^*$, where C^* acts on G as follows: it leaves the elements of T and of $Cv_2 + Cv_3$ fixed, and it acts by scalar multiplication on the 1-dimensional vector group spanned by v_1. We shall show that *(H, W) satisfies conditions (2) and (3) of Theorem 7.1, so that $W\backslash H$ has a quasi-affine structure. Thus we have a tower $W \subset G \subset H$ such that $W\backslash H$ has a quasi-affine structure, but $W\backslash G$ does not.* Again, this is in contrast with the situation for algebraic homogeneous spaces.

Clearly, we have $N_H(W) = V$, so that (3) is satisfied. With α denoting again the adjoint representation, we have $\alpha(H)^* = \alpha(V)\alpha(T \times C^*)^*$ and $\alpha(H)^*_V = \alpha(T \times C^*)^*_V$, which consists of the automorphisms $\rho_{(a,b,c)}$, where a, b, c are complex numbers, $ab \neq 0$, and the images of v_1, v_2, v_3 under $\rho_{(a,b,c)}$ are, respectively, av_1, $bv_2 + cv_3$, bv_3. The stabilizer F of W in $\alpha(T \times C^*)^*$ induces the automorphisms $\rho_{(a,a,0)}$ on V. If $(t, x) \in T \times C^*$ then the images of v_1, v_2, v_3 under the automorphism $\rho_{(a,a,0)}\alpha(t, x)$ are, respectively, $ae^{-t}xv_1$, $ae^t(v_2 + tv_3)$, ae^tv_3. Hence we see that $(F\alpha(T \times C^*))_V = \alpha(H)^*_V$, and hence that $S(W)\alpha(H) = \alpha(H)^*$, so that (2) is satisfied.

Finally, let us enlarge G by adjoining another normal vector component $U = Cu_1 + Cu_2 + Cu_3$ as a direct summand to V, and letting T act on U as follows: $t \cdot u_1 = e^{-ut}u_1$, $t \cdot u_2 = e^{ut}(u_2 + utu_3)$, $t \cdot u_3 = e^{ut}u_3$, where u is a fixed *irrational* number. Let G_0 denote the resulting enlarged group. We consider the following three vector subgroups of G_0:

$$W_0 = C(v_1 + v_2) + C(u_1 + u_2)$$
$$W_0^+ = C(v_1 + v_2) + Cv_3 + C(u_1 + u_2) + Cu_3$$
$$B = Cv_1 + Cv_2 + Cu_1 + Cu_2.$$

One sees as before that (G_0, W_0) violates (2), and that (G_0, W_0^+) satisfies (2). However, because of the irrationality of u, we have $N_{G_0}(W_0^+) = U + V$, so that (G_0, W_0^+) also satisfies (3). Hence $W_0^+\backslash G_0$ has a quasi-affine structure. It is not difficult to verify that B satisfies both (2) and (3), so that $B\backslash G_0$ has

a quasi-affine structure. On the other hand, we have evidently $W_0^+ \cap B = W_0$, and the subgroup generated by W_0^+ and B is the normal closed vector subgroup $U + V$ of G_0. Thus, in an independent generic notation, *there is a complex analytic group G with closed complex analytic subgroups X and Y (actually, vector groups) such that XY is a closed normal vector subgroup of G, both $X\backslash G$ and $Y\backslash G$ have quasi-affine structures, but $(X \cap Y)\backslash G$ does not have a quasi-affine structure.* This is again in contrast with what happens for algebraic groups [1, Th. 11], and it shows that Proposition 7.2 cannot be strengthened.

The lack of any transitivity properties with respect to the existence of quasi-affine structures that is illustrated by the above examples accounts for the complications in the proof of Theorem 7.1.

UNIVERSITY OF CALIFORNIA
AND
YALE UNIVERSITY.

REFERENCES.

[1] A. Bialynicki-Birula, G. Hochschild and G. D. Mostow, "Extensions of representations of algebraic linear groups," *American Journal of Mathematics*, vol. 85 (1963), pp. 131-144.

[2] C. Chevalley, *Fondements de la géométrie algébrique*, Faculté des Sciences de Paris, 1958 (mimeographed).

[3] ———, *Classification des groupes de Lie algébriques*, Séminaire, Ecole Normale Supérieur, 1956-1958, Exposé 8 (mimeographed).

[4] G. Hochschild, "Cohomology of algebraic linear groups," *Illinois Journal of Mathematics*, vol. 5 (1961), pp. 492-519.

[5] G. Hochschild and B. Kostant, "Differential forms and Lie algebra cohomology for algebraic linear groups," *Illinois Journal of Mathematics*, vol. 6 (1962), pp. 264-281.

[6] G. Hochschild and G. D. Mostow, "Representations and representative functions of Lie groups," *Annals of Mathematics*, vol. 66 (1957), pp. 495-542.

[7] ———, "Representations and representative functions of Lie groups," III, *Annals of Mathematics* (2), vol. 70 (1959), pp. 85-100.

[8] ———, "On the algebra of representative functions of an analytic group," *American Journal of Mathematics*, vol. 83 (1961), pp. 111-136.

[9] ———, "On the algebra of representative functions of an analytic group," II, *American Journal of Mathematics*, vol. 86 (1964), pp. 869-887.

[10] G. D. Mostow, "Self-adjoint groups," *Annals of Mathematics*, vol. 62 (1955), pp. 44 55.

A CERTAIN SUBGROUP OF THE FUNDAMENTAL GROUP.

By D. H. Gottlieb.

Introduction.[1] Let X be a topological space with x_0 as a base point A homotopy $H: X \times I \to X$ is called a *cyclic homotopy* if

$$H(x,0) = H(x,1) = x.$$

In another notation, h_t is a cyclic homotopy if $h_0 = h_1 = 1_X$, where 1_X denotes the identity map of X.

If h_t is a cyclic homotopy, the path given by $\sigma: I \to X$ such that $\sigma(s) = h_s(x_0)$ will be called the *trace* of h_t. The trace is obviously a closed path.

The set of homotopy classes of those loops which are the trace of some cyclic homotopy form a subgroup of the fundamental group which we shall denote by $G(X, x_0)$. It is the purpose of this paper to study $G(X, x_0)$, establish some elementary properties, compute it for one dimensional graphs, two dimensional compact manifolds, lens spaces and projective spaces. In addition its effects on the universal covering space and the mapping space X^X will be discussed.

Jaing Bo-Ju, in a recent paper [1], has also investigated this group. He was mostly interested in the role the group played in the Nielsen-Wecken theory of fixed point classes. Some properties of $G(X, x_0)$ proved here were mentioned by Jaing Bo-Ju, but they were not of the same generality except in the cases of Theorem 1. 8 and Theorem II. 4.

The present paper is divided into four parts. The first part deals with the elementary properties of $G(X, x_0)$, and $G(X, x_0)$ is computed for many kinds of spaces. In particular, Corollary I. 13 tells us that if X is aspherical, then $G(X, x_0)$ is the center of $\pi_1(X, x_0)$.

In the second section, the role of $G(X, x_0)$ as the subgroup of the group of deck transformations of the universial covering space is discussed, leading to the calculation of $G(X, x_0)$ for lens spaces and for projective spaces. Theorem II. 7 gives a condition for a homeomorphism to be in the center of a discrete group of homeomorphisms acting freely on a contractible space.

Received October 27, 1964.

[1] This work was supported by N. S. F. Grant 1908.

In part III, the relation of $G(X, x_0)$ to the mapping space X^X is discussed. If X is aspherical, it is shown that the identity component of X^X has the homotopy groups, $\pi_1(X^X, 1_X) \cong Z(\pi_1(X, x_0))$, the center of $\pi_1(X, x_0)$, and $\pi_n(X^X, 1_X) = 0$ for $n > 1$. M.-E. Hamstrom has investigated the homotopy groups of the space of homeomorphisms of two dimensional aspherical manifolds. They turn out to be the same as the homotopy groups of the space of mappings. This indicates a deeper relation between the mapping space and the subspace of homeomorphisms of manifolds.

In part IV, Theorem IV. 1 says that $G(X, x_0) = 1$ if X is compact and the Euler Poincaré number is zero. This fact may be of some use in computing Homeotopy groups; see G. S. McCarty [5]. Also we have Corollary IV. 3 which says that if $\chi(X) \neq 0$ and X is aspherical, then $Z(\pi_1(X)) = 1$. This result is applied to subcomplexes X of S^n and yields facts about $S^n - X$.

I. The group $G(X, x_0)$.

§ 1. $G(X, x_0)$.

We shall concern ourselves only with pathwise connected $C. W.$-complexes in this paper. Let X be one such with x_0 as a base point. We begin our investigation by inquiring; which loops are the trace of some cyclic homotopy? The first theorem shows that the answer depends only on the homotopy classes of the loops.

If σ is a loop. i. e., $\sigma: I \to X$ such that $\sigma(0) = \sigma(1) = x_0$, then $[\sigma]$ shall denote the equivalence class of all loops α homotopic to σ under a homotopy h_t such that $h_t(0) = h_t(1) = x_0$. In symbols, this will be written $\alpha \cong \sigma \operatorname{rel} x_0$. We shall also regard σ as a map from the circle (S^1, s_0) to (X, x_0) and $[\sigma]$ will denote the set of all α such that $\alpha \cong \sigma \operatorname{rel} x_0$.

THEOREM I. 1. *If σ is the trace of a cyclic homotopy, and $\alpha \in [\sigma]$, then α is the trace of a cyclic homotopy.*

Proof. Let $H: X \times I \to X$ be a cyclic homotopy with σ as its trace and let h_t be the homotopy connecting σ with α. Let L be the subcomplex of $X \times I$ given by $(X \times 0) \cup (X \times 1) \cup (x_0 \times I)$. Define a partial homotopy of H on L as follows: $k_s: L \to X$ such that $k_s(x, t) = x$ if $t = 0$ or $t = 1$ and $k_s(x_0, t) = h_s(t)$.

Now L is a sub complex of $X \times I$, and hence has the homotopy extension property. This means that there is a homotopy $K_t: X \times I \to X$ such that $K_0 = H$ and $K_t \mid L = k_t$. Then $K_1: X \times I \to X$ is a cyclic homotopy on X with trace α.

Definition. Let $G(X, x_0)$ be the set of all elements $[\sigma] \in \pi_1(X, x_0)$ such that σ is the trace of a cyclic homotopy on X.

THEOREM I. 2. *$G(X, x_0)$ is a subgroup of $\pi_1(X, x_0)$.*

Proof. Let $[\alpha]$ and $[\beta] \in G(X, x_0)$. Let h_t and k_t be the required cyclic homotopies respectively. Define a homotopy $l_t : X \to X$ such that $l_t(x) = h_{2t}(x)$ for $0 \leqq t \leqq \frac{1}{2}$ and $l_t(x) = k_{2t-1}(x)$ for $\frac{1}{2} \leqq t \leqq 1$. The trace of l_t is the loop $\alpha \cdot \beta$. Hence $[\alpha \cdot \beta] = [\alpha] \cdot [\beta] \in G(X, x_0)$.

Also $[\alpha]^{-1} \in G(X, x_0)$ since $[\alpha]^{-1} = [\alpha^{-1}]$ and α^{-1} is the trace of h_{1-t}: $X \to X$.

The next theorem shows that $G(X, x_0)$, viewed as a subgroup of $\pi_1(X, x_0)$ is independent of the choice of the base point x_0. Because of this, we shall abbreviate $G(X, x_0)$ to $G(X)$ when no confusion occurs.

Let $\sigma : I \to X$ be a path such that $\sigma(0) = x_0$ and $\sigma(1) = x_1 \in X$. Then σ induces an isomorphism $\sigma_* : \pi_1(X, x_1) \cong \pi_1(X, x_0)$ such that $\sigma_*([\alpha]) = [\sigma \cdot \alpha \cdot \sigma^{-1}]$.

THEOREM I. 3. $\sigma_* : G(X, x_1) \cong G(X, x_0)$.

Proof. Since σ_* is 1-1, all we must show is that $\sigma_*(G(X, x_1)) \subseteq G(X, x_0)$.

Let $[\alpha] \in G(X, x_1)$. Then there exists a cyclic homotopy $H : X \times I \to X$ with trace α. By the homotopy extension property, there is homotopy $J : X \times \mathrm{I} \to X$ such that $J(x, 0) = x$ and $J(x_0, t) = \sigma(t)$.

Define $K : X \times I \to X$ by

$$\begin{aligned} K(x,t) &= J(x, 3t), && 0 \leqq t \leqq \tfrac{1}{3} \\ K(x,t) &= H(J(x,1), 3t-1), && \tfrac{1}{3} \leqq t \leqq \tfrac{2}{3} \\ K(x,t) &= J(x, 3(1-t)), && \tfrac{2}{3} \leqq t \leqq 1. \end{aligned}$$

Now K is a cyclic homotopy and its trace with respect to x_0 is $\sigma \cdot \sigma \cdot \sigma^{-1}$. So $\sigma_*[\alpha] = [\sigma \cdot \alpha \cdot \alpha \sigma^{-1}] \in G(X, x_0)$.

§ 2. $P(X, x_0)$ And Computations.

We now establish some notation. Suppose (X, x_0) and (Y, y_0) are two spaces with base points, then we will always assume that $X \times Y$ has the base point (x_0, y_0). Also, X will denote $X \times y_0$ and Y will denote $x_0 \times Y$ and $X \vee Y = (X \times y_0) \cup (x_0 \times Y)$.

Remark I. Let $\sigma : (S^1, s_0) \to (X, x_0)$. Then $[\sigma] \in G(X, x_0)$ if and only if the map $f : X \vee S^1 \to X$ such that $f(x) = x$ whenever $x \in X$ and $f(s) = \sigma(s)$ if $s \in S^1$ can be extended to $X \times S^1$.

The elements of $\pi_1(X, x_0)$ operate on $\pi_n(X, x_0)$ as a group of automorphisms in a standard way, [4].

Definition. The set of elements of $\pi_1(X, x_0)$ which operate trivially on all $\pi_n(X, x_0)$ form a subgroup which will be denoted as $P(X, x_0)$.

Remark II. $[\alpha] \in \pi_1(X, x_0)$ operates trivially on $\pi_n(X, x_0)$ if and only if for every map $f\colon S^n \to X$, there exists an extension $F\colon S^n \times S^1 \to X$ such that $F \mid S^1 = \alpha$.

Now we are in position to prove the next theorem, whose corollaries will give us $G(X)$ for many spaces.

THEOREM I. 4. $G(X, x_0) \subseteq P(X, x_0)$.

Proof. Let $[\alpha] \in G(x, x_0)$. By Remark I, we have a map $H: X \times S^1 \to X$ such that $H \mid X = 1_x$ and $H \mid S^1 = \alpha$. Let $f\colon (S^n, r_0) \to (X, x_0)$ be any map from an n-sphere to X. We shall define a map $F\colon S^n \times S^1 \to X$ as follows; $F(r, s) = H(f(r), s)$ for $r \in S^n$ and $s \in S^1$. Since $F(r, s_0) = H(f(r), s_0) = f(r)$, $F \mid S^n = f$. Also $F(r_0, s) = H(x_0, s) = \alpha(s)$ implies that $F \mid S^1 = \alpha$. Therefore, by Remark II, $[\alpha] \in P(X, x_0)$.

The subgroup of $\pi_1(X, x_0)$ which operates trivially on $\pi_1(X, x_0)$ itself is precisely the center of $\pi_1(X)$, hereafter denoted by $Z(\pi_1(X))$. Thus $P(X, x_0) \subseteq Z(\pi_1(x))$ so we have $G(X) \subseteq Z(\pi_1(X))$.

COROLLARY I. 5. *If T is any 1-dimensional polyhedron except for the homotopy circle, then $G(T) = 1$.*

COROLLARY I. 6. *Let P^n be the projective space of dimension n. Then $G(P^{2n}) = 1$ for $n > 0$.*

Proof. P^{2n} is not a $2n$-simple space as is well known. That is $\pi_1(P^{2n})$ does not act trivially on $\pi_{2n}(P^{2n})$. Since $\pi_1(P^{2n}) \cong Z_2$, this means that the generator, α, of $\pi_1(P^{2n})$ does not act trivially on $\pi_{2n}(P^{2n})$. Hence $\alpha \notin P(P^{2n}, x_0)$, so $P(X, x_0)$ is the trivial subgroup. Thus $G(P^{2n})$ is trivial.

COROLLARY I. 7. *If M is any closed 2-dimensional manifold with the exception of the torus and the Klein, then $G(M) = 1$.*

Proof. If $M = P^2$, $G(M) = 1$ by the preceding corollary. Otherwise $\pi_1(M)$ has a trivial center as is well known. Hence $G(M) = 1$.

For two of the exceptional cases to Corollaries 5 and 7, the circle S^1 and the torus T, we see that $\pi_1(S^1) = G(S^1)$ and $\pi_1(T) = G(T)$. This result follows from the fact S^1 and T are both topological groups and the following theorem.

THEOREM I. 8. *If X is an H-space, then $G(X)=\pi_1(X)$.*

Proof. An H-space (X,l) has a continuous multiplication and an element l such that right and left multiplication are both homotopic to the identity on X. Since we are assuming that X is a $C.W.$ complex, $X \vee X$ is a subcomplex of $X \times X$ and so has the homotopy extension property. Hence there exists a continuous multiplication, $\cdot$, such that l is a multiplicative identity.

Let $\sigma: I \to X$ be any closed path in X such that $\sigma(0)=\sigma(1)=l$.

Define a cyclic homotopy as follows; $h_t(x)=\sigma(t)\cdot x$. The trace $\tau(t)=h_t(l)=\sigma(t)\cdot l=\sigma(t)$. Thus every closed loop at the identity is the trace of some cyclic homotopy, hence $G(X)=\pi_1(X)$.

§ 3. Properties of $G(X,x_0)$.

Any map $f:(X,x_0)\to(Y,y_0)$ induces a homomorphism $f_*:\pi_1(X,x_0)\to\pi_1(Y,y_0)$. This important property is not enjoyed by $G(X,x_0)$. That is $f_*(G(X))$ is not necessarily contained in $G(Y)$. This may be seen as follows.

Let Y be the figure eight. Let $i:(S^1,s_0)\to(Y,y_0)$ be the embedding of the circle onto one of the loops of Y. Now let α be a generator of $\pi_1(S^1,s_0)$. Then $i_*(\alpha)$ is not equal to the identity 1, of $\pi_1(Y,y_0)$. Since $G(Y,y_0)=1$, $i_*(\alpha)\notin G(Y,y_0)$. On the other hand, since $G(S^1)=\pi_1(S^1)$, $\alpha\in G(S^1,s_0)$. Thus $i_*(G(S^1))\not\subset G(Y)$.

All is not lost, for we do get the following theorems.

THEOREM I. 9. *Let $r:(X,x_0)\to(Y,y_0)$ be a retraction. Then*

$$r_*(G(X,x_0))\subseteq G(Y,y_0).$$

Proof. Let $i: Y\to X$ be the inclusion map. Let $y_0\in Y$ be the base point of Y and $i(y_0)$ be the base point of X. Let $[\alpha]\in G(X,i(y_0))$. Then there exists a map $K: X\times S^1\to X$ such that $K \mid X=1_X$ and $K\mid S^1=\alpha$.

Define a map $H: Y\times S^1\to Y$ by setting $H(y,s)=r\circ K(i(y),s)$ for $y\in Y$ and $s\in S^1$. Now $H(y,s_0)=r\circ K(i(y),s_0)=r(i(y))=y$ and $H(y_0,s)=r\circ K(i(y_0),s)=r(\alpha(s))=r\circ\alpha(s)$. Hence $[r\circ\alpha]\in G(Y,y_0)$. But $r_*[\alpha]=[r\circ\alpha]$, so $r_*(G(X,i(y_0))\subseteq G(Y,y_0)$.

Now consider any x_0 such that $r(x_0)=y_0$. Let σ be a path such that $\sigma(0)=i(y_0)$ and $\sigma(1)=x_0$. Then σ induces an isomorphism $\sigma_*:\pi_1(X,x_0)\cong\pi_1(X,i(y_0))$ as follows

$$\sigma_*[\alpha]=[\sigma\cdot\alpha\cdot\sigma^{-1}].$$

Let $[\alpha] \in G(X, x_0)$. By Theorem I. 3, $\sigma_*[\alpha] \in G(X, i(y_0))$ and so $r_*(\sigma_*[\alpha]) \in G(Y, y_0)$. But $r_*(\sigma_*[\alpha]) = r_*[\sigma \cdot \alpha \cdot \sigma^{-1}] = [r \circ \sigma \cdot r \circ \alpha \cdot r \circ \alpha^{-1}]$. Since $r \circ \sigma$ and $r \circ \sigma^{-1}$ are closed paths in Y, we have

$$r_*(\sigma_*[\alpha]) = [r \circ \sigma] \cdot [r \circ \alpha] \cdot [r \circ \sigma^{-1}] = [r \circ \sigma] \cdot [r \circ \sigma] \cdot [r \circ \sigma]^{-1}.$$

Since $r_*(\sigma_*[\alpha]) \in G(Y, y_0) \subseteq Z(\pi_1(Y, y_0))$, we have

$$\begin{aligned} r_*(\sigma_*[\alpha]) &= [r \circ \alpha]^{-1} \cdot r_*(\sigma_*[\alpha]) \cdot [r \circ \alpha] \\ &= [r \circ \alpha]^{-1} \cdot ([r \circ \sigma] \cdot [r \circ \alpha] \cdot [r \circ \sigma]^{-1}) \cdot [r \circ \sigma] = [r \circ \alpha]. \end{aligned}$$

Therefore $r_*[\alpha] \in G(Y, y_0)$.

If $f: (X, x_0) \to (Y, y_0)$ is a homotopy equivalence, then f induces an isomorphism $f_*: G(X, x_0) \to G(Y, y_0)$. This results from the following theorem.

THEOREM I. 10. *If $f: (X, x_0) \to (Y, y_0)$ is a homotopy equivalence, then $f_*(G(X, x_0)) = G(Y, y_0)$.*

Proof. That $f: X \to Y$ is a homotopy equivalence implies the existence, by definition, of a map $g: Y \to X$ such that $f \circ g \simeq 1_Y$ and $g \circ f \simeq 1_X$. Since y_0 has the Homotopy Extension Property in Y, we may assume that $g(y_0) = x_0$. Let $J: Y \times I \to Y$ be a homotopy such that $J(y, 0) = f \circ g(y)$ and $J(y, 1) = y$.

Now let $[\sigma] \in G(X, x_0)$. Since $f_*: \pi_1(X, x_0) \cong \pi_1(Y, y_0)$ is an isomorphism, we need merely to show that $f_*[\sigma] = [f \circ \sigma] \in G(Y, y_0)$. Let $h_t: X \to X$ be a cyclic homotopy such that $h_t(x_0) = \sigma(t)$ for all $t \in I$. Define a homotopy

$$K: Y \times I \to Y \text{ by } K(y, t) = (f \circ h_t \circ g)(y).$$

Then $k(y, 0) = f \circ g(y) = K(y, 1)$ and

$$k(y_0, t) = f(h_t(x_0)) = f(\sigma(t)) = f \circ \sigma(t)$$

for all $t \in I$.

Define $T: Y \times I \to Y$ such that

$$\begin{aligned} T(y, t) &= J(y, 1 - 3t) \text{ for } 0 \leqq t \leqq \tfrac{1}{3} \\ T(y, 3t - 1) &= K(y, 3t - 1) \text{ for } \tfrac{1}{3} \leqq t \leqq \tfrac{2}{3} \\ T(y, 3t - 2) &= J(y, 3t - 2) \text{ for } \tfrac{2}{3} \leqq t \leqq 1. \end{aligned}$$

Now $T(y, 0) = J(y, 1) = y$ and $T(y, 1) = J(y, 1) = y$, so T is a cyclic homotopy.

Let $\alpha\colon I \to Y$ be the path given by $\alpha(t) = J(y_0, t)$. Now $\alpha(0) = J(y_0, 0) = f \circ g(y_0) = f(x_0) = y_0$, and $\alpha(1) = J(y_0, 1) = y_0$, so α is a closed path. So the trace of T at y_0, τ, is given by:

$$\tau(t) = T(y_0, t) = (\alpha^{-1} \cdot (f \circ \sigma) \cdot \alpha)(t).$$

Hence $[\tau] = [\alpha]^{-1} \cdot [f \circ \sigma] \cdot [\alpha] \in G(Y, y_0)$. Hence $[f \circ \sigma] \in G(Y, y_0)$ since $G(Y) \subseteq Z(\pi_1(Y))$.

Another property G shares in common with the fundamental group is the following.

THEOREM I.11. $G(X \times Y, (x_0, y_0)) \cong G(X, x_0) \oplus G(Y, y_0)$.

Proof. Let $Z = X \times Y$ and $Z_0 = (X, y_0)$. There exists an isomorphism

$$h\colon \pi_1(Z, Z_0) \to \pi_1(X, y_0) \oplus \pi_1(Y, y_0),$$

such that

$$h([\alpha]) = p_*([\alpha]) \oplus q_*([\alpha])$$

where p_* and q_* are induced homomorphisms from the projections of Z onto X and Y respectively. Now $h(G(Z)) \subseteq G(X, x_0) \oplus G(Y, y_0)$ as may readily be seen by noting that projections are retractions and applying Theorem I.9 to the definition of h.

On the other hand, let $[\alpha]$ and $[\beta]$ be elements of $G(X, x_0)$ and $G(Y, y_0)$ respectively.

Now $h^{-1}([\alpha] \oplus [\beta]) = [(j \circ \alpha) \cdot (k \circ \beta)]$ where j and k inject $X \to X \times y_0$ and $Y \to x_0 \times Y$ respectively.

Since $[\alpha]$ and $[\beta]$ are in $G(X, x_0)$ and $G(Y, y_0)$ respectively, there exists cyclic homotopies H and J having traces α and β respectively.

Let $K\colon X \times Y \times I \to X \times Y$ be defined as follows:

$$K(x, y, t) = (H(x, 2t), y) \text{ for } 0 \leqq t \leqq \tfrac{1}{2}$$
$$K(x, y, t) = (x, J(y, 2 - 2t)) \text{ for } \tfrac{1}{2} \leqq t \leqq 1.$$

It can easily be verified that K is a cyclic homotopy on $X \times Y$ with trace $(j \circ \alpha) \cdot (k \circ \beta)$. Hence $h^{-1}([\alpha] \oplus [\beta]) \in G(Z, z_0)$, so $h^{-1}(G(X) \oplus G(Y)) \subseteq G(Z)$. Hence $h(G(Z)) \supseteq G(X) \oplus G(Y)$.

§ 4. Aspherical Spaces.

The fact that $G(X) \subseteq P(X)$ leads naturally to the questions; Is there a space X for which $G(X) \neq P(X)$, and if so, under what conditions does equality obtain? The author has not been able to answer these questions, but they have stimulated the next important theorem.

For every $[\sigma] \in \pi_1(X, x_0)$, we can define $f_\sigma\colon X \vee S^1 \to X$ such that $f_\sigma \mid X = 1_X$ and $f_\sigma \mid S^1 = \sigma$. Let $f_\sigma^{(n+1)}\colon (X \vee S^1) \cup (X^{(n)} \times S^1) \to X$ be an extension of f_σ where $X^{(n)}$ is the n-skeleton of X. If $f_\sigma^{(n+1)}$ exists, we say that $[\sigma]$ is $(n+1)$-*extensible*. The set of all $(n+1)$-extensible $[\sigma]$ forms a subgroup of $\pi_1(X_1, x_0)$ which we shall denote by $G^{(n)}(X, x_0)$. We get a descending sequence of groups as follows;

$$G^{(1)}(X) \supseteq G^{(2)}(X) \supseteq \cdots \supseteq G(X).$$

On the other hand, let $P^{(n)}(X, x_0)$ stand for the subgroup of $\pi_1(X, x_0)$ of all $[\sigma]$ which operate trivially on $\pi_k(X, x_0)$ for $k \leqq n$. Then

$$P^{(1)}(X) \supseteq P^{(2)}(X) \supseteq \cdots \supseteq P(X).$$

THEOREM I. 12. $G^{(1)}(X, x_0) = P^{(1)}(X, x_0) = Z(\pi_1(X, x_0))$.

Proof. $P^{(1)}(X, x_0) = Z(\pi_1(X, x_0))$ is well known. To prove $G^{(1)}(X, x_0) = Z(\pi_1(X, x_0))$, we must show that $f_\sigma^{(2)}$ exists iff $[\sigma] \in Z(\pi_1(X_1, x_0))$. That is, we must show that $f_\sigma\colon X \vee S^1 \to X$ is 2-extensible over $X \times S^1$ iff $[\sigma] \in Z(\pi_1(X, x_0))$. This shall be shown by appealing to the following result, which can be found in [4], p. 194.

Remark III. Let L be a connected subcomplex of K containing v_0 and $f\colon (L, v_0) \to (Y, y_0)$ be a map into a pathwise connected space Y. Then f and the inclusion map $i\colon L \subset K$ induce the homeomorphisms

$$f_*\colon \pi_1(L, v_0) \to \pi_1(Y, y_0), \quad i_*\colon \pi_1(L, v_0) \to \pi_1(K, v_0).$$

Then f is 2-extensible over K iff there exists a homomorphism $h\colon \pi_1(K, v_0) \to \pi_1(Y, y_0)$ such that $f_* = hi_*$.

For the case at hand, let $L = X \vee S^1$, $Y = X$, $K = X \times S^1$ and $f = f_\sigma$. Then $\pi_1(L) \cong \pi(X) * \pi(S^1)$, the free product, and $\pi_1(K) \cong \pi_1(X) \oplus \pi_1(S^1)$.

Now let $\alpha \in \pi_1(X)$ and $\beta \in \pi_1(S^1)$. Let multiplication between elements of $\pi_1(X)$ be expressed by $(\cdot)$. Let ν generate $\pi_1(S^1)$. Then $i_*(\alpha * \beta) = \alpha \oplus \beta$ and $f_*(\alpha * \nu) = \alpha \cdot [\sigma]$.

Suppose that h exists such that $f_* = hi_*$. Now $f_*(\nu) = [\sigma]$ and $i_*(\nu) = 1 \oplus \nu$, so $[\sigma] = h(1 \oplus \nu)$. On the other hand, $\alpha = f_*(\alpha) = hi_*(\alpha) = h(\alpha \oplus 1)$.

Thus h, if it exists, must satisfy the equation $h(\alpha \oplus \nu) = \alpha \cdot [\sigma]$ for all $\alpha \in \pi_1(X)$. Now $f_*(\alpha * \nu) = \alpha \cdot [\sigma]$ and $f_*(\nu * \alpha) = [\sigma] \cdot \alpha$. But $hi_*(\alpha * \nu) = \alpha \cdot [\sigma] = hi_*(\nu * \alpha)$. Hence $\alpha \cdot [\sigma] = [\sigma] \cdot \alpha$ for all $\alpha \in \pi_1(X)$.

The above theorem enables us to determine precisely what $G(X)$ is when X is aspherical, i. e., when $\pi_n(X) = 0$ for $n > 1$.

COROLLARY I. 13. *If X is aspherical, then $G(X, x_0) = Z(\pi_1(X, x_0))$.*

Proof. Since X is aspherical, $X \times S^1$ is aspherical. Thus any map $f_\sigma\colon X \vee S^1 \to X$ which is 2-extensible must be extensible over $X \times S^1$.

The above corollary permits us to settle the one holdout among the closed, 2-dimensional manifolds the Klein Bottle K.

COROLLARY I. 14. *Let K be the Klein Bottle. Then $G(K) = Z(\pi_1(K))$.*

II. The universal covering space. As in the first chapter, X will always be a pathwise-connected $C.W.$-complex. This is enough to insure the existence of the universal covering space C. We shall let $p\colon (C, \tilde{x}_0) \to (X, x_0)$ be the covering projection.

§1. *The Universal Covering Space and $G(X, x_0)$.*

There is a natural isomorphism, ν, between $\pi_1(X, x_0)$ and the group of Deck Transformations, $\mathcal{D}(X)$, acting on C. Thus $G(X, x_0)$ corresponds to a subgroup of $\mathcal{D}(X)$ under ν. This subgroup, $\nu G(X)$, has a natural definition within $\mathcal{D}(X)$.

THEOREM II. 1. *$G(X, x_0)$ is isomorphic to the subgroup of those Deck Transformations which are homotopic to 1_C by fiber preserving homotopies.*

Proof. Suppose that $[\lambda] \in \pi_1(X, x_0)$ gives rise to the deck-transformation $l\colon C \to C$. This means that any path α from $\tilde{x}_0$ to $l(\tilde{x}_0)$ projects down upon the closed path $p \circ \alpha \in [\lambda]$.

Now suppose that $[\lambda] \in G(X, x_0)$. Then there exists a cyclic homotopy $h_t\colon X \to X$ whose trace is λ. Now $1_C\colon C \to C$ covers the map $1_X \circ p\colon C \to X$. Since $h_t \circ p\colon C \to X$ is a homotopy of $1_X \circ p$, by the Covering Homotopy Property there must exist a homotopy $\tilde{h}_t\colon C \to C$ which lifts $h_t \circ p$. That is $h_t \circ p = p \circ \tilde{h}_t$. Now $h_1 = 1_X$ so $p = p \circ \tilde{h}_1$. Thus $\tilde{h}_1$ must be a deck transformation of C. Now $\tilde{h}_1 = l$ since the path $\tilde{\tau}(t) = \tilde{h}_t(\tilde{x}_0)$ running from $\tilde{x}_0$ to $\tilde{h}_1(\tilde{x}_0)$ lifts λ. So $\tilde{h}_t$ is the required fiber preserving homotopy from 1_C to l.

Conversely, if $\tilde{h}_t$ is a fiber preserving homotopy such that $\tilde{h}_0 = 1_C$ and $\tilde{h}_1 = l$, then there exists a cyclic homotopy $h_t\colon X \to X$ such that $h_t \circ p = p \circ \tilde{h}_t$. Clearly h_t is a cyclic homotopy and its trace $\tau(t) = h_t(x_0)$ is contained in $[\lambda]$.

For covering spaces, fiber-preserving homotopies satisfy a very nice condition.

THEOREM II. 2. *The homotopy $\tilde{h}_t\colon C \to C$ is fiber preserving iff $f \circ \tilde{h}_t = \tilde{h}_t \circ f$ for every $f \in \mathcal{D}(X)$.*

Proof. Suppose $\tilde{h}_t$ is fiber preserving. Let f be any deck transformation

and $x \in C$ any point. Then $f(x)$ and x are in the same fiber. Since $\bar{h}_t(x)$ and $\bar{h}_t(f(x))$ are both in the same fiber, there is a $g \in \mathfrak{D}(X)$ such that $g \circ \bar{h}_t(x) = \bar{h}_t \circ f(x)$. If $\epsilon > 0$ is sufficiently small, $g \circ \bar{h}_{t-\epsilon}(x) = \bar{h}_{t-\epsilon} \circ f(x)$. Thus the greatest lower bound of the set of t's such that $g \circ \bar{h}_t(x) = \bar{h}_t \circ f(x)$ must occur when $t = 0$. Therefore by continuity, $g \circ \bar{h}_0(x) = \bar{h}_0 \circ f(x)$. But $\bar{h}_0 = 1_C$, so $g(x) = f(x)$. This can occur only when $g = f$. Thus $f \circ \bar{h}_t = \bar{h}_t \circ f$ for all $f \in \mathfrak{D}(X)$.

On the other hand, suppose $f \circ \bar{h}_t = \bar{h}_t \circ f$ for all $f \in \mathfrak{D}(X)$. Let x and y both be in the same fiber of p and suppose that $f(x) = y$. Now $\bar{h}_t = f^{-1} \circ \bar{h}_t \circ f$, so $\bar{h}_t(x) = f^{-1} \circ \bar{h}_t(y)$. Thus $\bar{h}_t(x)$ is in the same fiber as $\bar{h}_t(y)$. Hence $\bar{h}_t$ is fiber preserving.

COROLLARY II.3. *$G(X, x_0)$ is isomorphic to the subgroup of $\mathfrak{D}(X)$ given by those deck transformations which are homotopic to the identity by a homotopy which commutes with every deck transformation.*

§ 2. Computations.

Let p and q be relatively prime integers. Then $L(p, q)$, a three dimensional lens space, has a fundamental group isomorphic to the cyclic group of order P.

THEOREM II.4. $G(L(p,q)) = \pi_1(L(p,q))$.

Proof. Let S^3 be the 3-sphere given by the complex coordinates (Z_0, Z_1) such that $Z_0\bar{Z}_0 + Z_1\bar{Z}_1 = 1$. Then let $f\colon S^3 \to S^3$ such that

$$f(Z_0, Z_1) = (Z_0 e^{2\pi i/p}, Z_1 e^{2\pi q i/p}).$$

Now f generates a cyclic group Z_p of rotations, each element of which is fixed point free. The factor space S^3/Z_p is the lens space $L(p, q)$. [3, page 262]

Now let $h_t\colon S^3 \to S^3$ such that $h_t(Z_0, Z_1) = (Z_0 e^{2\pi t i/p}, Z_1 e^{2\pi q t i/p})$ be a homotopy. Now h_0 is the identity on S^3 and $h_1 = f$. Also $f \circ h_t = h_t \circ f$, so h_t commutes with all Z_p. Therefore $f \in \nu G(L(p,q))$, hence $G(L(p,q)) \cong Z_p$.

THEOREM II.5. *Let P^n be the real projective space of dimension n. Then $G(P^{2n+1}) = \pi_1(P^{2n+1}) \cong Z_2$.*

Proof. S^{2n+1} can be given by the $n+1$ tuple of complex numbers $(Z_0, \cdots, Z_n)$ satisfying the equation $\bar{Z}_0 Z_0 + \cdots + \bar{Z}_n Z_n = 1$. The projective space P^{2n+1} is created by identifying antipodal points. Let f be the deck transformation such that $f(Z_0, \cdots, Z_n) = (-Z_0, \cdots, -Z_n)$. Define a homotopy $h_t\colon S^{2n+1} \to S^{2n+1}$ such that $h_t(Z_0, \cdots, Z_n) = Z_0 e^{\pi t i}, \cdots, Z_n e^{\pi t i})$.

Then h_0 is the identity and $h_1 = f$. Also $f \circ h_t = h_t \circ f$. So by Corollary II. 3, $G(P^{2n+1}) = \pi_1(P^{2n+1}) \cong Z_2$.

§ 3. $\mathfrak{A}(X)$.

Definition. Let $\mathfrak{A}(X)$ be the set of all those deck transformations in the center of $\mathfrak{D}(X)$ which are homotopic to the identity. It is easy to verify that $\mathfrak{A}(X)$ is a subgroup of $\mathfrak{D}(X)$. We shall use the same symbol, $\mathfrak{A}(X)$, to stand for the corresponding subgroup in $\pi_1(X, x_0)$.

THEOREM II. 6. $G(X, x_0) \subseteq \mathfrak{A}(X) \subseteq P(X, x_0)$.

Proof. That $G(X, x_0) \subseteq \mathfrak{A}(X)$ is obvious from Corollary II. 3.

Let $f \in \mathfrak{D}(X)$ such that $f \cong 1_X$. Let $h_t \colon C \to C$ be the homotopy such that $h_0 = 1_X$ and $h_1 = f$. Let $\tilde{\phi} \colon I \to C$ such that $\tilde{\phi}(t) = h_t(x_0)$. Let $\phi = p \circ \tilde{\phi}$. Then f corresponds to $[\phi]$ under ν.

Suppose ϕ operates on $[\alpha] \in \pi_n(X, x_0)$, $n > 1$. Then ϕ operates trivially on α iff the map $g \colon S^n \vee S^1 \to X$ such that $g \mid S^n = \alpha$ and $g \mid S^1 = \phi$ can be extended to a map $g' \colon S^n \times S^1 \to X$.

We define g' as follows. There exists a map $l \colon S^n \to C$ such that $p \circ l = g \mid S^n$ for $n > 1$. This is a well known property of the universal covering space. If we consider S^1 as the unit interval such that 0 and 1 are identified, then $g'(s, t) = p \circ h_t \circ l, s)$. Since $g'(s, 0) = P(l(s)) = p \circ f(l(s)) = p \circ h_1 \circ l(s) = g'(s, 1)$, g' is well defined and it is easily verified that g' is an extension of g. So we have shown that $\mathfrak{A}(X)$ is contained in the subgroup of all elements of $\pi_1(X, x_0)$ which operates trivially on $\pi_n(X, x_0)$ for $n > 1$. Since $\mathfrak{A}(X) \subseteq Z(\pi_1(X, x_0))$, $\phi \in \mathfrak{A}(X)$ operates trivially on $\pi_1(X, x_0)$, hence $\mathfrak{A}(X) \subseteq P(X, x_0)$.

We have no examples which indicate whether the inclusions are proper. Certainly, for spaces whose universal covering spaces are compact, odd-dimensional homotopy spheres, $\mathfrak{A}(X) = P(X) = Z(\pi_1(X))$. In particular, if X is a compact three-dimensional manifold with finite fundamental group, then $\mathfrak{A}(X) = Z(\pi_1(X))$.

§ 4. Aspherical Spaces.

THEOREM II. 7. *Let X be a contractible C. W.-complex with π, a discrete group of homeomorphisms of X onto itself, acting freely on X. Then if $f \in Z(\pi)$, there is a homotopy h_t which commutes with g for all $g \in \pi$ such that $h_0 = 1_X$ and $h_1 = f$.*

Proof. If we let X/π stand for the space obtained by identifying the

orbits under π, then X may be regarded as the universal covering space of X/π. Thus π may be regarded as the deck transformations of the covering and hence also as the fundamental group of X/π. Since X/π is aspherical, $G(X/\pi) \cong Z(\pi)$. Hence the center of π consists of homeomorphisms of X which are homotopic to the identity by a homotopy h_t such that $f \circ h_t = h_t \circ f$.

III. X^X.

Let X^X denote the space of continuous mappings from X into X with the compact-open topology. Let Ω be the path connected component of X^X which contains the identity 1_X.

Let $p\colon X^X \to X$ be the evaluation $p(f) = f(x_0)$. Since we wish p to be continuous, we will assume that X is locally compact throughout this chapter. We also avoid complications if p is a fibering, and this occurs when X is a locally finite simplicial polyhedron. With the help of p, we can characterize $G(X, x_0)$.

Remark IV. There is a natural homeomorphism between the space of maps $(X^X)^{S^n}$ and $X^{X \times S^n}$ given by $\phi\colon (X^X)^{S^n} \to X^{X \times S^n}$ such that $\phi(f)(x, s) = (f(s))(x)$ for $x \in X$ and $s \in S^n$. Note that $f \cong g$ iff $\phi(f) \cong \phi(g)$.

THEOREM III. 1. $p_* \pi_1(X^X, 1_X) = G(X, x_0)$.

Proof. By the remark, the closed path $f\colon S^1 \to X^X$ corresponds to the cyclic homotopy $\phi(f)\colon X \times S^1 \to X$. Now $p \circ f\colon S^1 \to X$ is equal to $\phi(f) \mid S^1$ for $p(f)(x_0, s) = f(s)(x_0) = p(f(s)) = p \circ f(s)$.

This is to say that every closed loop f in $\Omega \subseteq X^X$ is a cyclic homotopy of X whose trace equals $p \circ f$ and conversely, every cyclic homotopy of X is a closed path f in Ω such that $p \circ f$ equals the trace of the cyclic homotopy.

THEOREM III. 2. *Let X be a locally finite, aspherical, pathwise connected simplicial polyhedron. Then $p_*\colon \pi_1(X^X, 1_X) \cong Z(\pi_1(X, x_0))$ and $\pi_n(X^X, 1_X) = 0$ for $n > 1$.*

Proof.

LEMMA 1. $\pi_n(X^X, 1_X) = 0$ *if* $n > 1$.

Proof. Let $f\colon X \times S^n \to X$ such that $f(X, s_0) = x$ for all $x \in X$ where $s_0 \in S^n$ is the base point of S^n. Define $d\colon X \times S^n \to X$ such that $d(x, s) = x$, that is d is the projection of $X \times S^1$ onto X. By the remark, if we can show that $f \cong d$, then $\phi^{-1}(f) \cong \phi^{-1}(d)$. Since $\phi^{-1}(d)\colon S^1 \to X^X$ is the constant map onto 1_X, this will prove the lemma.

Since X is aspherical, $f \simeq d$ iff $f_*: \pi_1(X \times S^n, x_0 \times s_0) \to \pi_1(X, x_0)$ and $d_*: \pi_1(X \times S^n, x_0 \times s_0) \to \pi_1(X, x_0)$ are equivalent, that is

$$f_*(\alpha) = \xi^{-1} \cdot g_*(\alpha) \cdot \xi$$

for all $\alpha \in \pi_1(X \times S^n)$ and some $\xi \in \pi_1(X)$. See Hu [4], pp. 198-9.

Now $\pi_1(X \times S^n, x_0 \times s_0) \cong \pi_1(X, x_0) \oplus \pi_1(S^n, s_0) \cong \pi_1(X, x_0)$. Both f_* and d_* "act like the identity" and so $f_* = d_*$. Hence $f \simeq d$.

LEMMA 2. $p_*(\pi_1(X^X, 1_X)) = Z(X, x_0))$.

Proof. Since X is aspherical, $Z(\pi_1(X, x_0)) = G(X, x_0)$. Hence by the preceding theorem, the lemma is true.

LEMMA 3. *Let $\Omega_0 \subseteq X^X$ be the space of maps such that $f(x_0) = x_0$ for all $f \in \Omega_0$. Then $\pi_1(\Omega_0, 1_X) = 0$.*

Proof. Let $d: X \times S^1 \to X$ such that $d(x, s) = x$. Let f be any arbitrary $f: X \times S^1 \to X$ such that $f(x_0, s) = x_0$ for all $s \in S^1$. We will prove the lemma by showing there is a homotopy $h_t: X \times S^1 \to X$ such that $h_0 = f$ and $h_1 = d$ and $h_t(x_0, s) = x_0$ for all $t \in I$ and $s \in S^1$. For then $\phi^{-1}(h_t)$ will be a homotopy connecting $\phi^{-1}(f) \in \Omega_0$ and $\phi^{-1}(d)$ which is the constant map $S^1 \to 1_X$. Since $\phi^{-1}(h_t) \in \Omega_0$ for each $t \in I$, the lemma will be proved.

We may regard S^1 as I with the points 0 and 1 identified. Thus we may regard f and d as maps from $X \times I$ into X.

Let

$$\begin{aligned} A = (X \times 0 \times I) \cup (X \times 1 \times I) \cup (x_0 \times I \times I) \\ \cup (X \times I \times 0) \cup (X \times I \times 1). \end{aligned}$$

Define $H^{(1)}: A \to X$ such that

$$\begin{aligned} H^{(1)}(x, s, 0) &= f(x, s) \\ H^{(1)}(x, s, 1) &= d(x, s) \\ H^{(1)}(x_0, s, t) &= x_0 \\ H^{(1)}(x, 0, t) &= H(x, 1, t) = x. \end{aligned}$$

We wish to extend $H^{(1)}$ to a map $H: X \times I \times I \to X$. Then $H(x, s, t) = h_t(x, s)$ will give us the homotopy mentioned above, which will prove the lemma.

Let $X^{(n)}$ be the n-skeleton of X. Let $K = X \times I \times I$. Regard I as being decomposed into $\{0\}$, $\{1\}$ and $(0, 1)$. Then

$$K^{(1)} \subseteq A \text{ and } K^{(2)} \subseteq X^{(0)} \times I \times I \cup A.$$

We shall extend $H^{(1)}: A \to X$ to $H^{(2)}: K^{(2)} \to X$ by the following procedure. Let $x_i \in X^{(0)}$. Then

$$S_i^1 = (x_i \times I \times 0) \cup (x_i \times 1 \times I) \cup (x_i \times I \times 1) \cup (x_i \times 0 \times I)$$

forms a circle. Since $S_i^1 \in A$, $H^{(1)} \mid S_i^1 : S_i^1 \to X$. It is easily seen, that $H^{(1)} \mid S_i^1$ is null homotopic and hence may be extended to $H_i^{(2)}: X_i \times I \times I$.

Define $H^{(2)}: K^{(2)} \to X$ by

$$H^{(2)}(y) = H^{(1)}(y) \text{ if } y \in A$$
$$H^{(2)}(y) = H_i^{(2)}(y) \text{ if } y \in x_i \times I \times I.$$

Since X is aspherical, we may extend $H^{(2)}: K^{(2)} \to X$ to $H: X \times I \times I \to X$. Since $H(x, 0, t) = H(x, 1, t)$, H may be regarded as a map from $X \times S^1 \times I$ to X. Now we can define $h_t(x, s) = H(x, s, t)$ and we see that $h_0 = f$ and $h_1 = d$ and $h_t(x_0, s) = x_0$.

LEMMA 4. $p_*: \pi_1(X^X, 1_X) \cong Z(\pi_1(X, x_0))$.

Proof. Consider the homotopy sequence

$$\pi_1(\Omega_0) \xrightarrow{i_*} \pi_1(X^X) \xrightarrow{p_*} \pi_1(X).$$

Since $\pi_1(\Omega_0) = 0$, p_* must be 1-1. But $p_* \pi_1(X^X) = Z(\pi_1(X))$.

Lemmas 1 through 4 prove the theorem.

COROLLARY III. 3. *If X is a pathwise connected aspherical locally finite simplicial polyhedron, then Ω, the path component of X^X containing 1_X, is contractible when $Z(\pi_1(x)) = 1$.*

Proof. By Milnor [6], Ω has the homotopy type of a $C.W.$-complex. Since $\pi_n(\Omega) = 0$ for all n, by a theorem of Whitehead's [8], Ω is contractible.

COROLLARY III. 4. *If X is a pathwise connected, aspherical, locally finite simplicial polyhedron, then $p: \Omega \to X$ is a homotopy equivalence iff $\pi_1(X, x_0)$ is abelian.*

Proof. Again by Milnor [6] and Whitehead [8].

IV. The Euler-Poincaré number and $G(X)$.

THEOREM IV. 1. *Suppose X has the same homotopy type as a compact, connected polyhedron. Then if the Euler-Poincaré number $\chi(X)$ is not equal to zero, $G(X)$ is trivial.*

Proof. By Theorem I. 10, we may assume that X is a compact, connected polyhedron.

The proof is a simple application of the Nielsen-Wecken theory of fixed point classes. We shall summarize the pertinent facts needed for the proof. These are proved in Wecken [7] and are in the notation of Jaing Bo-Ju in [1].

Let $\tilde{X}$ be the universal covering of X. We regard $\pi_1(X)$ as the group of deck transformations on $\tilde{X}$. Let $f: X \to X$. Consider the set of all lifts of f to maps $\tilde{f}: \tilde{X} \to \tilde{X}$. We define an equivalence relation among these lifts as follows: $\tilde{f} \equiv \tilde{f}_1$ if and only if $\tilde{f}_1 = \gamma^{-1} \circ \tilde{f} \circ \gamma$ for some $\gamma \in \pi_1(X)$. Let $[\tilde{f}]$ denote the equivalence class of $\tilde{f}$. The set of fixed points of $\tilde{f}$ project down, by the covering map p, onto a subset of fixed points of $\tilde{f}$. The fixed points of any $\tilde{f}_1$ in the same equivalence class as $\tilde{f}$ also project down to the same subset of fixed points of f. If $\tilde{f}_1$ is not equivalent to f, then the fixed points of $\tilde{f}_1$ project down to a subset of fixed points of f disjoint from those of $\tilde{f}$. This procedure partitions the fixed points of f into disjoint subsets, called fixed point classes. Thus each fixed point class is uniquely associated with an equivalence class of lifts of f. We can also have lifts, $\tilde{f}$, of f with no fixed points, and so the equivalence class of $\tilde{f}$ corresponds to a void class of fixed points.

If $h_t: f \cong g$ for $g: X \to X$, then h_t defines a 1-1 correspondence between the lifts of f and those of g preserving equivalence classes. Hence there is a 1-1 correspondence between fixed point classes.

With each fixed point class $[\tilde{f}]$, it is possible to assign a number ν such that $\nu = 0$ if $[\tilde{f}]$ is empty and such that ν is preserved under homotopy. That is if $[\tilde{f}]$ corresponds to $[\tilde{g}]$ under a homotopy from f to g, then ν for $[\tilde{g}]$ is equal to the ν for $[\tilde{f}]$. Finally the sum of all the ν's equals Λ_f, the Lefschitz number.

Suppose that $f = 1_X$. Then every $\nu = 0$ except possibly for ν_1, the number associated with the fixed point class given by the identity $\tilde{1}: \tilde{X} \to \tilde{X}$. This follows since every other lift of 1_X has no fixed point. Also we know that $\Lambda_f = \chi(X)$ when $f = 1_X$. Assume that $\chi(X) \neq 0$. Then $\nu_1 = \chi(X) \neq 0$.

Let $\alpha \in G(X)$. Then there is a cyclic homotopy $h_t: X \to X$ which can be lifted to a homotopy $\tilde{h}_t: \tilde{1} \cong \alpha$ where we regard α as a deck transformation. So $[\tilde{1}]$ corresponds to $[\alpha]$. But $\alpha: \tilde{X} \to \tilde{X}$ has no fixed points, unless $\alpha = \tilde{1}$. Since $\nu \neq 0$ for $[\alpha]$, the associated fixed point class must be nonempty so $\alpha = \tilde{1}$. Thus $\alpha = 1 \in \pi_1(X)$. Hence $G(X) = 1$.

This theorem yields a number of very interesting corollaries.

COROLLARY IV. 2. *Let X be the homotopy type of a connected, compact polyhedron. If X is an H-space and $\chi(X) \neq 0$, then $\pi_1(X) = 1$.*

Proof. By Theorem I. 8, $G(X) = \pi_1(X)$. Hence, since $G(X) = 1$, we have $\pi_1(X) = 1$.

As a matter of fact, it can be shown, using homological properties of H-spaces, that $\chi(X) = 0$ or $\chi(X) = 1$, in which case X is contractible. See [2] for a proof of this in the case of semigroups.

COROLLARY IV. 3. *Let X have the same homotopy type as a connected, compact polyhedron. If $\chi(X) \neq 0$ and X is aspherical, then $Z(\pi_1(X)) = 1$.*

Proof. By Corollary I. 13, $G(X) = Z(\pi_1(X))$. Hence $Z(\pi_1(X)) = 0$.

As an application of this result, we can get the following well known result.

COROLLARY IV. 4. *For any closed 2-dimensional manifold, excepting the torus, projective space and the Klein Bottle, the center of the fundamental group is trivial.*

Corollary IV. 3 also has applications to the imbedding of complexes in spheres. The author is indebted to L. P. Neuwirth for suggesting this line of approach.

COROLLARY IV. 5. *Let X be an m-dimensional, connected subcomplex of S^n where $m \leqq n-2$. Then $S^n - X$ aspherical implies that $Z(\pi_1(S^n - X)) = 1$ provided that $\chi(X) \neq 0$ if n is odd and $\chi(X) \neq 2$ if n is even.*

Proof. Let $\pi^i(Y)$ stand for the i-th Betti number of any topological space Y for $i > 0$. For $i = 0$, $\pi^i(Y)$ will equal the number of connected components of Y minus 1. Now $\pi^p(X) = \pi^{n-p-1}(S^n - X)$ for $0 \leqq p \leqq n-1$ by Alexander's Duality. Then

$$\begin{aligned} \chi(X) &= \sum_{p=0}^{m} (-1)^p \pi^p(X) + 1 \\ &= \sum_{p=0}^{m} (-1)^p \pi^{n-p-1}(S^n - X) + 1 \\ &= \sum_{j=n-1}^{n-m-1} (-1)^{n-j-1} \pi^j (S^n - X) + 1 \\ &= \sum_{j=0}^{n-1} (-1)^{n-j-1} \pi^j (S^n - X) + 1 \end{aligned}$$

since $\pi^j(S^n - X) = \pi^{n-j-1}(X) = 0$ for $j < n - m - 1$. Also since $\pi^n(S^n - X) = 0$, we have

$$\chi(X)=\sum_{j=0}^{n}(-1)^{n-j-1}\pi^j(S^n-X)+1$$
$$=(-1)^{n-1}[\sum_{j=0}^{n}(-1)^j\pi^j(S^n-X)]+1$$
$$\chi(X)=(-1)^{n-1}[\chi(S^n-X)-1]+1$$
$$=-(-1)^n\chi(S^n-X)+(-1)^n+1.$$

Hence we have that

$$\chi(S^n-X)=(-1)^{n-1}\chi(X)+(1+(-1)^n).$$

So $\chi(S_n-X)\neq 0$, if $\chi(X)\neq 0$ when n is odd and also if $\chi(X)\neq 2$ if n is even.

Now S^n-X is connected and is of the same homotopy type as a closed subcomplex of S^n. Hence apply Corollary IV. 3.

A natural generalization of Theorem IV. 1 is the following: If X is a compact polyhedron and $\chi(X)\neq 0$, then $p_*\pi_n(X^X,1_X)=0$ for all n.

This statement is untrue. It is known that the homeotopy sequence [5] gives rise to isomorphisms $p_*:\pi_n(G)\cong\pi_n(S^2)$, $n>2$ where G is the group of homeomorphisms of S^2 onto itself and p_* is induced by the evaluation map. Since $p_*\pi_n(X^X,1_X)\supset p_*\pi_n(G,1_X)$, we see that $p_*\pi_3(X^X,1_X)=\pi_3(S^2)\cong Z$ if $X=S^2$. $\chi(S^2)\neq 0$ so the above generalization is false.

INSTITUTE FOR DEFENSE ANALYSES,
PRINCETON, NEW JERSEY.
UNIVERSITY OF ILLINOIS.

REFERENCES.

[1] Jaing Bo-Ju, "Estimation of the Nielsen Numbers," *Chinese Mathematics*, vol. 5 Issue #2, 1964.

[2] D. H. Gottlieb and N. J. Rothman, "Contractibility of certain semigroups," *Bulletin of the American Mathematical Society*, November 1964.

[3] Hilton and Wylie, *Homology Theory*, Cambridge University Press, 1960.

[4] S. T. Hu, *Homotopy Theory*, Academic Press, 1959.

[5] G. S. McCarty, "Homotopy theory," *Transactions of the American Mathematical Society*, vol. 106 (1963), pp. 293-304.

[6] J. W. Milnor, "On spaces having the same homotopy type as C. W. complexes," *Transactions of the American Mathematical Society*, vol. 90 (1959), pp. 272-280.

[7] F. Wecken, "Fixpunktklassen, I-III," *Mathematische Annalen*, vol. 117 (1940-1941); 118 (1941-1942).

[8] J. H. C. Whitehead, "Combinatorial homotopy I," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 213-245.

SUMS OF SQUARES AND CUSP FORMS.

By R. A. Rankin.

It is well known that $r_s(n)$, the number of representations of a positive integer n as the sum of s squares, can be expressed in the form

$$r_s(n) = \rho_s(n) + R_s(n),$$

where $\rho_s(n)$ is a 'divisor function' and $R_s(n)$ is the n-th Fourier coefficient of a cusp form; see [2], for example. When $s \leqq 8$, $R_s(n) = 0$ for all positive integers n.

In the language of modular form theory, this may be expressed by stating that

$$\vartheta_3{}^s(z) = E_s(z) + f_s(z), \tag{1}$$

where $E_s(z)$ is the Eisenstein series associated with the cusp ∞, $f_s(z)$ is a cusp form, and f_s vanishes identically for $s = 5$, 6, 7 and 8. Here we restrict our attention to $s > 4$ to avoid convergence difficulties and write, as usual,

$$\vartheta_3(z) = \sum_{n=-\infty}^{\infty} e^{\pi i n^2 z}.$$

Then

$$E_s(z) = \sum_{n=0}^{\infty} \rho_s(n) e^{\pi i n z} \quad \text{and} \quad f_s(z) = \sum_{n=1}^{\infty} R_s(n) e^{\pi i n z}$$

for $z \in \mathfrak{H} = \{z : \operatorname{Im} z > 0\}$.

At a recent conference in Madison, Wisconsin, the question was raised whether there exist any integers $s > 8$ for which $f_s = 0$. It is the purpose of this note to answer this question in the negative. In fact, rather more is proved, namely that, if s is any *real* number greater than 8, then the cusp form f_s in (1) does not vanish identically. (For $4 < s \leqq 8$, f_s does vanish identically, since the zero form is the only cusp form in these cases.)

Since ϑ_3 has no zeros in $\mathfrak{H}$, it is possible to define $\vartheta_3{}^s$, for every real s, as a holomorphic function on $\mathfrak{H}$ with a Fourier series

$$\vartheta_3{}^s(z) = \sum_{n=0}^{\infty} r_s(n) e^{\pi i n z}$$

Received October 29, 1964.

in which $r_s(0) = 1$. Further, $\vartheta_3{}^s \in \mathfrak{M}_s = \{\Gamma, -\frac{1}{2}s, v_s\}$, where this denotes the vector space of all holomorphic modular forms of dimension $-\frac{1}{2}s$ for the group Γ, with multiplier system v_s. Here Γ is the subgroup of index 3 in the modular group $\Gamma(1)$ consisting of all matrices

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

for which $T \equiv I \pmod 2$ or $T \equiv V \pmod 2$, where

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad V = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix};$$

we also put

$$U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad L = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

The multiplier system v_s is defined on the generators U^2 and V of Γ by

$$v_s(U^2) = 1, \qquad v_s(V) = e^{-\pi i s/4}.$$

The dimension of the vector space $\mathfrak{M}_s$ is $1 + [s/8]$ for $s \geqq 0$, where square brackets denote the integral part.

As a fundamental region for Γ we take the set defined by

$$z \in \mathfrak{H}, \qquad |z + \tfrac{1}{2}| \leqq \tfrac{1}{2}, \qquad -1 \leqq \operatorname{Re} z \leqq 0.$$

This has three cusps, namely ∞, $0 = V\infty$ and $-1 = L\infty$, the first two being congruent modulo Γ.

For $s > 4$, $E_s \in \mathfrak{M}_s$, where E_s is the Eisenstein series defined by

$$E_s(z) = \sum \{v_s(T)(cz + d)^{s/2}\}^{-1},$$

the summation being over a maximal set of matrices $T \in \Gamma$ having different second rows $[c, d]$ and $[-c, -d]$ does not occur if $[c, d]$ does. Then (1) holds for some cusp form $f_s \in \mathfrak{M}_s$, when $s > 4$.

If $g \in \mathfrak{M}_s$, the order of g at the cusp -1 is defined to be the order of the transformed function

$$g_L(z) = g(z) \mid L = z^{-s/2} g(Lz)$$

at ∞, and for this purpose we need to examine its Fourier series, which will be of the form

$$g_L(z) = \sum_{m+\kappa \geqq 0} a_m e^{2\pi i (m+\kappa) z}$$

where $\kappa = \{s/8\}$, curly brackets denoting the fractional part. Then $\operatorname{ord}(g, -1, \Gamma)$ is defined to be the least value of $m + \kappa$ such that $a_m \neq 0$.

Since $\vartheta_3^s(z) \mid L = e^{-\pi is/4}\vartheta_2^s(z)$, where

$$\vartheta_2(z) = 2\sum_{n=1}^{\infty} e^{\pi i(n-\frac{1}{2})^2 z},$$

it follows that, for $s \geqq 0$,

$$\operatorname{ord}(\vartheta_3^s, -1, \Gamma) = s/8.$$

On the other hand, we shall show that, for $s > 4$,

$$\operatorname{ord}(E_s, -1, \Gamma) = \{s/8\}^* = 1 - \{-s/8\}. \tag{2}$$

Since $\{s/8\}^* \neq s/8$, for $s > 8$, it follows that $\vartheta_3^s \neq E_s$ for $s > 8$, as claimed. Because $E_s = \vartheta_3^s$ for $4 < s \leqq 8$, we need only prove (2) for $s > 8$.

The Fourier series for $E_s \mid L$ can be found from formula (14) of a paper by H. Petersson [1]. We apply this result with $r = \frac{1}{2}s$, $\Gamma_0 = L^{-1}\Gamma L$ and $A = L$, so that

$$E_s(z) \mid L = \tfrac{1}{2}E_{-r}(z; v; A, \Gamma_0),$$

where v is the conjugate multiplier system for the group Γ_0. Petersson's parameters take the following values:

$$e_0 = 2,\ \delta = 0,\quad N_0 = 1,\quad \eta = \kappa = \{s/8\}.$$

We therefore have, writing $N = \{s/8\}^* - \kappa$,

$$E_s(z) \mid L = \frac{(2\pi)^{s/2}e^{-\pi is/4}}{\Gamma(s/2)}\sum_{n=N}^{\infty}(n+\kappa)^{\frac{1}{2}s-1}a_n e^{2\pi i(n+\kappa)z},$$

where

$$a_n = \sum_{c>0}\frac{\sigma_c}{c^{s/2}}\sum_{d}\frac{e^{2\pi i(n+\kappa)d/c}}{\lambda(c,d)},$$

and is real, as it is not hard to show. Here σ_c and $\lambda(c,d)$ are numbers of unit modulus. Also $0 \leqq d < c$, $(c,d) = 1$ and only values of c and d occur for which, for some corresponding matrix $T \in \Gamma(1)$, we have $T \in L\Gamma_0 = \Gamma L$. It follows that c must be odd and so, since $c = 1$, $d = 0$ is a possible choice $(T = L)$ and $\sigma_1 = \lambda(1,0) = 1$, we have, for $n \geqq N$,

$$a_n \geqq 1 - \sum_{\substack{c \geqq 3 \\ c \text{ odd}}}\phi(c)c^{-s/2}$$

$$= 2 - \frac{\zeta(\frac{1}{2}s-1)(1-2^{1-\frac{1}{2}s})}{\zeta(\frac{1}{2}s)(1-2^{-\frac{1}{2}s})}. \tag{3}$$

It is easily verified that the right-hand side of (3) is positive for $s > 8$

(in fact, it is positive for $s > 4.622$), and so, in particular, $a_N \neq 0$, from which (2) follows.

In conclusion, we note that an alternative method of proof is to show that

$$2s = r_s(1) > \rho_s(1). \tag{4}$$

For this purpose, $\rho_s(1)$ can be calculated from Petersson's formula in a similar way, giving

$$|\rho_s(n)| \leqq \frac{\pi^r n^{r-1} \zeta(r-1)}{(1-2^{-r})\zeta(r)\Gamma(r)} = c_s n^{r-1} \qquad (r = \tfrac{1}{2}s) \tag{5}$$

for $n \geqq 1$. It may be verified that $c_s < 2s$ for $s > s_0$, where s_0 is approximately $8 \cdot 7$. The coefficient $\rho_s(n)$ is real and, taking $n = 1$, we deduce that (4) holds when $s > s_0$. This alternative method fails, however, when $8 < s < s_0$, since then $c_s > 2s$.

THE UNIVERSITY,
GLASGOW, SCOTLAND.

REFERENCES.

[1] H. Peterson, "Über die Entwicklungskoeffizienten der ganzen Modulformen und ihre Bedeutung für die Zahlentheorie, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 8 (1931), pp. 215-242.

[2] R. A. Rankin, "On the representation of a number as the sum of any number of squares, and in particular of twenty," *Acta Arithmetica*, vol. 7 (1962), pp. 399-407.

AN INFINITE DIMENSIONAL VERSION OF SARD'S THEOREM.

By S. Smale.*

The purpose of this note is to introduce a non-linear version of Fredholm operators and to prove that in this context, Sard's Theorem holds if zero measure is replaced by first category (Section 1). We give applications to local uniqueness in non-linear elliptic equations (Section 2) and to define a notion of degree (Section 3).

Section 1. We recall first some definitions and facts from the linear theory (see [3], [4]) for details. A *Fredholm operator* is a continuous linear map $L: E_1 \to E_2$ from one Banach Space to another with the properties:

a) $\dim \operatorname{Ker} L < \infty$

b) Range L is closed

c) $\operatorname{Coker} L = E_{2/\text{Range}\, L}$ has finite dimension.

If L is Fredholm, then its *index* is $\dim \operatorname{Ker} L - \dim \operatorname{Coker} L$ so that the index of L is an integer.

(1.1) Theorem. *The set $F(E_1, E_2)$ of Fredholm operators is open in the space of all bounded operators $L(E_1 E_2)$ in the norm topology. Furthermore the index is continuous on $F(E_1, E_2)$.*

For a proof see [4].

The non-linear extension of the preceding notion seems to fit best into the context of differentiable manifolds locally like Banach spaces (see Lang [5]), called here differentiable manifolds and denoted by M, V. We will assume all our manifolds to be connected and to have a countable base.

A *Fredholm map* is a C' map $f: M \to V$ such that for each $x \in M$, the derivative $Df(x): T_x(M) \to T_{f(x)}(V)$ is a Fredholm operator. The *index* of f is defined to be the index of $Df(x)$ for some x. By (1.1), since f is C' and M is connected, the definition doesn't depend on x.

Let $f: M \to V$ be any C' map. A point $x \in M$ is called a *regular point* of f if $Df(x): T_x(M) \to T_{f(x)}(V)$ is surjective and is *singular* if not regular. The images of the singular points under f are called the *singular values* or

Received November 2, 1964.

* This research was partially supported by NONR 3656(14) and NSF GP 2497.

critical values and their complement the *regular values.* Note that if $y \in V$ is not in the image of f it is automatically a regular value. We will need the following.

(1.2) SARD THEOREM. *Let U be an open set of R^p and $f\colon U \to R^q$ be a C^s map where $s > \max(p-q, 0)$. Then the set of critical values in R^q has measure zero.*

For a proof see [8] or [9].

We will say *almost all* instead of "except for a set of first category" etc.

Our main theorem is

(1.3) THEOREM. *Let $f\colon M \to V$ be a C^q Fredholm map with $q > \max(\operatorname{index} f, 0)$. Then the regular values of f are almost all of V.*

The condition that f be Fredholm is necessary from the example of I. Kupka (to be published) of a C^∞ real-function on Hilbert space with critical values containing an open set. His example extends easily to give a C^∞ map from one separable Hilbert space to another with critical values possessing an interior point.

Strictly speaking (1.3) is a generalization of a theorem of A. B. Brown, an earlier special case of Sard's Theorem [A. B. Brown, *Transactions of the American Mathematical Society,* vol. 38 (1935), pp. 379-394].

(1.4) COROLLARY. *If $f\colon M \to V$ is a Fredholm map of negative index, its image contains no interior points.*

(1.5) COROLLARY. *If $f\colon M \to V$ is a C^q Fredholm map,* $q > \max$ $(\operatorname{Index} f, 0)$, *then for almost all $y \in V$, $f^{-1}(y)$ is a submanifold of M whose dimension is equal to index f or is empty.*

We now prove (1.3). Since M has a countable base and first category is closed under countable union, it is sufficient to prove the theorem locally. Thus we can assume given $f\colon U \to E'$, where U is an open set of a Banach space E, and E' is another Banach space.

Let $x_0 \in U$, $A = Df(x_0)\colon E \to E'$. Since $\dim \operatorname{Ker} A < \infty$, E can be written in the form $E_1 \times \operatorname{Ker} A$, E_1 a Banach space and $x_0 = (p_0, q_0) p_0 \in E_1$, $q_0 \in \operatorname{Ker} A$. Then the first partial derivative $D_1 f(p,q)\colon E_1 \to E$ maps E_1 injectively onto a closed subspace of E for all (p,q) sufficiently close to (p_0, q_0). Thus using the implicit function theorem, we can choose a product neighborhood $D_1 \times D_2$ of (p_0, q_0) in $E_1 \times \operatorname{Ker} A$ such that D_2 is compact and if $q \in D_2$, f restricted to $D_1 \times q$ is a (differentiable) homeomorphism onto its image.

(1.6) THEOREM. *A Fredholm map is locally proper. In other words,*

if $f: M \to V$ is Fredholm and $x \in M$, there exists a neighborhood N of x such that f restricted to N is proper.

A map is *proper* if the inverse image of a compact set is compact.

To prove (1.6) choose $N(x) = D_1 \times D_2$ as above and let $f(x_i) = y_i \to y$, $x_i = (p_i, q_i) \in D_1 \times D_2$. It is sufficient to show that the x_i have a convergent subsequence. Since D_2 is compact we may assume $q_i \to q$ and since $f(p_i, q) \to y$ even that $q_i = q$. But f restricted to $D_1 \times q$ is a homeomorphism onto its image, so $p_i \to p$, proving (1.6).

To prove (1.3), let $x_0 \in M$ and choose again $D_1 \times D_2 \subset E_1 \times \operatorname{Ker} A$ as above. The critical points of f are closed and therefore by the preceding theorem it is sufficient, given a neighborhood U_1 of $f(x_0)$ in E', to find a regular value of f in U_1.

Let $\pi: E' \to E'/\operatorname{Range} A$ be the projection. From the hypotheses of (1.3), (1.2) applies to the map $\phi: p_0 \times \operatorname{Ker} A \to E'/\operatorname{Range} A$ defined by $\phi(q) = \pi \cdot f(p_0, q)$ to give as a regular value z of ϕ in πU_1. Let $y \in \pi^{-1}(z) \cap U_1$. Then y is our desired regular value and our proof is finished.

Section 2. We will prove a local uniqueness theorem for the case of non-linear elliptic equations of 2nd order for domain in R^n with Dirichlet boundary conditions. Obviously the proof is valid in much greater generality.

Let Ω be a bounded domain in R^n with boundary, $\partial\bar{\Omega}$ a smooth sub-manifold. Define $C^k(\bar{\Omega})$, k a non-negative integer to be the Banach space of C^k functions on $\bar{\Omega}$ with the C^k norm (see [1] or [2]). Let α satisfy $0 < \alpha < 1$, and let $C^{k+\alpha}(\bar{\Omega})$ be the space of C^k functions on $\bar{\Omega}$ with k-th derivative Hölder continuous, exponent α, endowed with the corresponding Banach space structure (again see [1] for example). If $f_0 \in C^{2+\alpha}(\bar{\Omega})$ let $C_{f_0}{}^{2+\alpha}(\bar{\Omega})$ denote the affine subspace of $C^{2+\alpha}(\bar{\Omega})$ of maps which agree with f_0 on $\partial\bar{\Omega}$.

Let $J^2(\bar{\Omega}) = \bar{\Omega} \times R \times L(R^n, R) \times L_s{}^2(R^n, R) = \{(x, p_0, p_1, p_2)\}$ where $L(R^n, R)$ denotes the space of linear maps $R^n \to R$, and $L_s{}^2(R^n, R)$ the space of symmetric bilinear maps $R^n \times R^n \to R$. Thus if $f \in C^2(\bar{\Omega})$ we can define $j_2 f: \bar{\Omega} \to J^2(\bar{\Omega})$ by $j_2(f)(x) = (x, f(x), D^1 f(x), D^2 f(x))$, D^1, D^2 denoting the first and second derivative respectively of f at x.

A (*non-linear*) *partial differential equation*, $\Phi(u) = g$ (2nd order etc.) is defined by a map $F: J^2(\bar{\Omega}) \to R$ which we will assume to be C^∞, as follows. Let $\Phi: S^s(\bar{\Omega}) \to C^{s-2}(\bar{\Omega})$ be the map $\Phi(u)(x) = F(j_2 u)(x)$, $s \geqq 2$. We will consider solutions of $\Phi(u) = f$ for given $f \in C^{s-2}(\bar{\Omega})$. Of course $\Phi u = g$ can be written as $F(x, u(x), Du(x), D^2 u(x)) = g(x)$.

We say that F (or Φ) is *elliptic* if the partial derivative

$$F_{p_2}(x, p_0, p_1, p_2) \in L_s{}^2(R^n; R)$$

is a positive or negative definite bilinear form for each $(x, p_0, p_1, p_2) \in J^2(\bar{\Omega})$.

(2.1) Suppose $F\colon J^2(\bar{\Omega})\ R$ is elliptic as above and *linear* (i.e., for each x is linear on $x \times R \times L(R^n, R) \times L_s^2(R^n, R)$), so $L = \Phi\colon C_{f_0}^{2+\alpha}(\bar{\Omega}) \to C^\alpha(\bar{\Omega})$ defined by F is also linear, $f_0 \equiv 0$. Then L is a Fredholm optrator of index zero.

This is well known and essentially proved in such greater generality in [1] and [2].

As an immediate consequence of (2.1) we obtain

(2.2) Lemma. *If $F\colon J^2(\bar{\Omega}) \to R$ is an elliptic non-linear partial differential equation and $f_0 \in C^{2+\alpha}(\bar{\Omega})$, then the induced map $\Phi\colon C_{f_0}^{2+\alpha}(\bar{\Omega}) \to C^\alpha(\bar{\Omega})$ is a Fredholm map of index zero with derivatives at $u \in C_{f_0}^{2+\alpha}(\bar{\Omega})$ given by*

$$D\Phi(u)(\eta) = \sum_{i=0}^{2} F_{p_i}(x, u(x), Du(x), D^2u(x)) D^i\eta(x)$$

where $\eta \in C_0^{2+\alpha}(\bar{\Omega})$ (i.e. $\eta \in C^{2+\alpha}(\bar{\Omega})$, $\eta(x) = 0$ for $x \in \partial\bar{\Omega}$), and $D^0\eta = \eta$.

From (1.5) and (2.2) we obtain "local uniqueness" for solutions of $\Phi(u) = g$ in the following form.

(2.3) Theorem. *If $F\colon J^2(\bar{\Omega}) \to R$ is an elliptic non-linear differential equation and $f_0 \in C^{2+\alpha}(\bar{\Omega})$, for almost all $g \in C^\alpha(\bar{\Omega})$, the set of $u \in C_{f_0}^{2+\alpha}(\bar{\Omega})$ such that $F(x, u(x), Du(x), D^2u(x)) = g(x)$ is discrete.*

Section 3. Let $f\colon M \to V$ be a C^r map and $g\colon W \to V$ be a C^r imbedding. We say that f is transversal to g if for each $(x,y) \in M \times W$ such that $f(x) = g(y)$ the spaces range $Df(x)$, range $Dg(y)$, span the tangent space $T_{f(x)}(V)$.

(3.1) Theorem. *Let $f\colon M \to V$, be a C^q Fredholm map and $g\colon W \to V$ a C^r imbedding of a finite dimensional manifold W with*

$$q > \max(\operatorname{Index} f + \dim W, 0).$$

Then there exists a C^r approximation g' of g such that f is transversal to g'. Furthermore if f is transversal to the restriction of g to a closed set A of W. then g' may be chosen so that $g' = g$ on A.

We remark that the usual finite dimensional version of this theorem [10] gives an approximation f' of f which requires a partition of unity on M, but requires less differentiability. In view of applications in the direction of Section 2, we wish to avoid such an assumption on M.

Since M and W have a countable base, a standard argument reduces the proof of (3.1) to a local lemma, namely the following.

(3.2) LEMMA. *Let $y \in W$, M, W as in* (3.1). *Then there exists a neighborhood U_1 of y and for any $\epsilon > 0$ an $\epsilon C'$ approximation g' of g such that f is transversal to g'/U_1.*

Proof of (3.2). We can assume that we have a neighborhood of $g(y)$ in V given as follows. $U_2 \subset R^p$, $V =$ Banach space $F = R^p \times F_1$ and $g\colon U_2 \to R^p \times 0$ is the identity $\times 0$, $\pi\colon F \to F_1$ the projection. Let U_1 be a neighborhood of y with $\mathrm{Cl}\, U_1 \subset \text{interior}\, U_2$ and ϕ a C^∞ function 1 on U_1, 0 on exterior U_2. Now by (1.3) let $z \in F_1$ be close to 0 and such that $\pi \cdot f$ has z as a regular value on $f^{-1}(g(U_2))$. Define g' as the translate of g by z on U_1 smoothed by ϕ.

This proves (3.2) and therefore (3.1) (That the last statement of (3.1) is true is clear from the argument).

(3.3) THEOREM. *Let $f\colon M \to V$ be a Fredholm map transversal to the imbedding $g\colon W \to V$, $\dim W < \infty$. Then $f^{-1}(g(w))$ is a submanifold of M, of dimension equal to the index $f + \dim W$.*

The proof is the same as the finite dimensional case [10], as well as for the modification of (3.3) to the case that $\partial W \neq \phi$.

We wish to define a generalized degree for proper Fredholm maps.

Recall that a *closed* map is one in which the image of a closed set is closed. Quite generally a proper map is closed and a closed map such that the inverse images of points are compact is proper.

A Fredholm map is locally proper (1.6) and the interested reader will be able to verify the following lemma.

(3.4) LEMMA. *If $f\colon M \to N$ is a Fredholm map and is closed where $\dim M = \infty$, then f is proper.*

Let η be the non-oriented cobordism ring of Thom (see [10]) with η^p consisting of classes of p dimensional manifolds, $p = 0, 1, \cdots$.

Let $f\colon M \to N$ (reminding the reader that M and N are connected) be a C^q proper Fredholm map, $q > p + 1$ where $\mathrm{p} = \operatorname{index} f \geqq 0$. We define an invariant (generalized degree mod 2) $\gamma(f) \in \eta^p$ as the class of $f^{-1}(y)$, y a regular value of f (see 1.5). To see that $\gamma(f)$ is independent of y let y_1 be another regular value of f and suppose $g\colon I \to V$ is an imbedding of the unit interval with $g(0) = y$, $g(1) = y_1$. By (3.1) we suppose f is transversal to g. Then by (3.3) $f^{-1}(g(I))$ effects a cobordism relation between $f^{-1}(y)$ and $f^{-1}(y_1)$, thus giving us the invariance of $\gamma(f)$. Remembering that a point of V not in the range of f is automatically a regular value, we summarize.

(3.5) THEOREM. *Let $f\colon M \to V$ be a C^q proper Fredholm map with*

$p = \operatorname{index} f \geqq 0$, *and* $q > p+1$. *Then there exists an invariant* $\gamma(f) \in \eta^p$ *defined by the non-oriented cobordism class of* $f^{-1}(y)$, y *some regular value of* f. *If* $\gamma(f) \neq 0$, f *is surjective.*

If the index of f is zero then $f^{-1}(y)$ is a finite set of points and $\gamma(f)$ is the ordinary degree reduced mod 2.

This is related to the Leray-Schauder degree [6], [7].

REFERENCES.

[1] Agmon, Douglis, Nirenberg, "Estimates for solutions of elliptic partial differential equations satisfying general boundary conditions, I," *Communications on Pure and Applied Mathematics*, vol. 12 (1959), pp. 623-727.

[2] F. Browder, "Functional analysis and partial differential equations, II," *Mathematische Annalen*, vol. 145 (1962), pp. 81-226.

[3] Cordes and Labrousse, "The invariance of the index in the metric space of closed operators," *Journal of Mathematics and Mechanics*, vol. 12 (1963), pp. 693-720.

[4] Gohberg and Krein, "The basic propositions on defect numbers and indices of linear operators," *Transactions of the American Mathematical Society*, vol. 13 (1960), pp. 185-264.

[5] S. Lang, *Introduction to differentiable manifolds*, New York, 1962.

[6] J. Leray, "La Thèorie des points fixes et ses applications en analyse," *Proceedings of the International Congress of Mathematicians* (1950), Providence, R. I.

[7] M. Nagumo, "Degree of mapping in convex linear topological vector spaces," *American Journal of Mathematics*, vol. 73 (1951), pp. 497-511.

[8] A. Sard, "The measure of the critical values of differentiable maps," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 883-890.

[9] S. Sternberg, *Lectures on Differential Geometry*, Englewood Cliffs, New Jersey, 1964.

[10] R. Thom, "Quelques propriétés globales des variétés différentiables," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 17-86.

ON THE PARTITIONS OF AN INTEGER INTO DISTINCT ODD SUMMANDS.

By Peter Hagis, Jr.

1. Introduction. In two recent papers ([2] and [3]) I have employed the Hardy-Ramanujan circle dissection method as refined by Rademacher [5] to obtain convergent series representations and asymptotic formulae for $q(n)$ and $Q(n)$, the number of partitions of a positive integer n into odd, and odd and unequal, parts respectively. The purpose of the present paper is to give an alternative formulation of the infinite series for $Q(n)$ which, while equivalent to that derived in [3], is much simpler in form. In the application of the circle dissection method there are two requirements. First, transformation equations must be found which exhibit the behavior of the generating function near its singularities at the rational points on the unit circle. Second, rather sharp estimates of the magnitude of certain complicated sums of roots of unity are needed. In [2] both the transformation equations and the estimates of the exponential sums were obtained by using procedures similar to those of Lehner [4]. In [3] analogous results were obtained from those in [2] by employing algebraic and number theoretic arguments. In the present paper our argument will be parallel to, but independent of, that in [3].

The key to this new attack is to replace the study of $Q(n)$ by that of $S(n)=(-1)^nQ(n)$. In Section 2 the generating function of $S(n)$ is given and the necessary transformation equations are obtained. In Section 3 the required exponential sum estimates are proved, and in Section 4 a convergent series for $S(n)$ is derived. As an immediate corollary we have a similar series for $Q(n)$ and in Section 5 this infinite series is compared with that obtained in [3]. Free use will be made throughout of the results and arguments to be found in [2].

2. The generating functions and transformation equations. The generating function for $q(n)$ is

$$F(x)=\prod_{m=0}^{\infty}(1-x^{2m+1})^{-1}=\sum_{n=0}^{\infty}q(n)x^n$$

Received October 15, 1964.
Revised January 7, 1965.

where $|x| < 1$ and where we have defined $q(0) = 1$. With $S(0) = 1$ we easily verify that the generating function for $S(n) = (-1)^n Q(n)$ is

$$H(x) = \prod_{m=0}^{\infty} (1 - x^{2m+1}) = 1/F(x) = \sum_{n=0}^{\infty} S(n) x^n$$

where $|x| < 1$.

For convenience we first give the transformation equation for $F(x)$, proved in [2]. In what follows k is a positive integer, $(h,k) = 1$, and $\mathcal{R}(z) > 0$. $hh' \equiv -1 \pmod{k}$, and $2hH' \equiv -1 \pmod{k}$ if k is odd.

THEOREM 1. *If k is even*

$$F(\exp\{2\pi ih/k - 2\pi z/k\}) = \omega(h,k)\exp\{\pi(z - 1/z)/12k\} \cdot F(\exp\{2\pi ih'/k - 2\pi/kz\}),$$

and if k is odd

$$F(\exp\{2\pi ih/k - 2\pi z/k\}) = 2^{-\frac{1}{2}}\chi(h,k)\exp\{\pi(2z + 1/z)/24k\} \cdot H(\exp\{2\pi iH'/k - \pi/kz\}).$$

Here,

$$(2.1) \qquad \begin{aligned} &\omega(h,k) = \exp\{\pi i\sigma(h,k)\}, \\ &\sigma(h,k) = \sum((\mu/k))((h\mu/k)), \ \mu = 1, 3, \cdots, k-1; \end{aligned}$$

$$(2.2) \qquad \begin{aligned} &\chi(h,k) = \exp\{\pi it(h,k)\}, \\ &t(h,k) = \sum((\mu/2k))((h\mu/k)), \mu = 1, 3, \cdots, 2k-1. \end{aligned}$$

$((y)) = y - [y] - \frac{1}{2}$ if y is not an integer, and $((y)) = 0$ if y is an integer.

Taking reciprocals in Theorem 1, we obtain the transformation equation for $H(x)$.

THEOREM 2. *If k is even*

$$H(\exp\{2\pi ih/k - 2\pi z/k\}) = \omega^*(h,k)\exp\{-\pi(z - 1/z)/12k\} \cdot H(\exp\{2\pi ih'/k - 2\pi/kz\}),$$

and if k is odd

$$H(\exp\{2\pi ih/k - 2\pi z/k\}) = 2^{\frac{1}{2}}\chi^*(h,k)\exp\{-\pi(2z + 1/z)/24k\} \cdot F(\exp\{2\pi iH'/k - \pi/kz\}).$$

Here $\omega^*(h,k) = 1/\omega(h,k)$ and $\chi^*(h,k) = 1/\chi(h,k)$.

We remark hat this result is much simpler than Theorem 1 of [3] which gives the transformation equation for

$$G(x) = \prod_{m=0}^{\infty} (1 + x^{2m+1}) = \sum_{n=0}^{\infty} Q(n) x^n,$$

the generating function of $Q(n)$.

3. Estimates of two exponential sums.

THEOREM 3. *If σ_1 and σ_2 are integers such that $0 \leqq \sigma_1 < \sigma_2 \leqq k$, and if, modulo k, one has $\sigma_1 \leqq h' < \sigma_2$, then the sums*

$$S_1 = \sum_{h \bmod k}{}' \omega^*(h,k)\exp\{-2\pi i(hn - h'\nu)/k\}$$

where $2 \mid k$, and

$$S_2 = \sum_{h \bmod k}{}' \chi^*(h,k)\exp\{-2\pi i(hn - H'\nu)/k\}$$

where $2 \nmid k$, are each subject to the estimate $O(n^{1/8}k^{2/3+\epsilon})$ uniformly in ν, σ_1, σ_2. $\sum'$ indicates that h runs over a reduced residue system modulo k.

Proof. The discussion in Sections 3, 4 and 5 of [2] holds word for word here if $\omega(h,k)$ and $\chi(h,k)$ are replaced by $\omega^*(h,k)$ and $\chi^*(h,k)$ throughout. The only differences occur in the algebraic signs of certain terms. Thus, if $4 \mid k$ we have

$$\omega^*(h,k) = \exp\{2\pi i(\phi(uh + vh')/Gk + 6\phi(c-2)/G)\}$$

where $f = 1$, $G = 48$ if $3 \mid k$, while $f = 3$, $G = 16$ if $3 \nmid k$. ϕ satisfies the congruence $f\phi \equiv 1 \pmod{Gk}$, $hh' \equiv -1 \pmod{Gk}$ and $u = 2 + 6k - 2k^2$, $v = -2 - k^2$. $c = 1, 3$ according as $h \equiv 1, 3 \pmod 4$.

If $k \equiv 2 \pmod 4$

$$\omega^*(h,k) = \exp\{2\pi i(\phi(h - h')/Gk^* - \Gamma A(d,k)/f)\}$$

where $f = 16$, $G = 3$ if $3 \mid k$, while $f = 48$, $G = 1$ if $3 \nmid k$. ϕ and Γ satisfy the congruences $f\phi \equiv 1 \pmod{Gk^*}$, $Gk^*\Gamma \equiv 1 \pmod f$ where $k = 2k^*$. $hh' \equiv -1 \pmod{Gk}$ and $A(d,k) = -12d(k + k^*) + 4d + 3(2k - k^2)$ where $d = -3, 3$ according as $h \equiv 1, 3 \pmod 4$. The apparent dependence of $A(d,k)$ on h is illusory since we easily verify that $A(3,k) \equiv A(-3,k) \pmod f$.

If $2 \nmid k$,

$$\chi^*(h,k) = \exp\{2\pi i(\Phi(2h + h')/gk - \gamma(12k(1-k) - 6(k\alpha - 1))/F)\}$$

where $F = 16$, $g = 3$ if $3 \mid k$, while $F = 48$, $g = 1$ if $3 \nmid k$. Φ, γ, α satisfy the congruences $F\Phi \equiv 1 \pmod{gk}$, $gk\gamma \equiv 1 \pmod F$, $k\alpha \equiv 1 \pmod 4$ with $0 < \alpha < 4$. $hh' \equiv -1 \pmod{gk}$.

Using these expressions the proof is completed by using the arguments in Section 7 of [1].

4. Convergent series for $S(n)$ and $Q(n)$. We are now prepared to find an exact formula for $S(n)$. By Cauchy's integral formula,

$$S(n) = \frac{1}{2\pi i}\int_C x^{-n-1}H(x)\,dx = \sum_{h,k}{}' \frac{1}{2\pi i}\int_{\xi_{h,k}} x^{-n-1}H(x)\,dx$$

where $0 \leqq h < k \leqq N$ and ξ_{hk} are the Farey arcs of order N of C, the circle $|x| = \exp\{-2\pi N^{-2}\}$. If on ξ_{hk} we let

$$x = \exp\{-2\pi N^{-2} + 2\pi ih/k + 2\pi i\phi\}$$

and write $w = N^{-2} - i\phi$, $z = wk$ we obtain

$$S(n) = \sum_{h,k}{}' \exp\{-2\pi inh/k\} \int H(\exp\{2\pi ih/k - 2\pi z/k\})\exp(2\pi nw)\,d\phi.$$

The integration, here and in the rest of this section, is from $-1/k(k+k_1)$ to $1/k(k+k_2)$ where k_1, k, k_2 are the denominators of consecutive terms in the Farey series of order N.

We now split the sum over k into two parts $S(n,0)$ and $S(n,1)$ according to whether k is even or odd respectively. By Theorem 2,

$$S(n,0) = \sum_{h,k}{}' \omega^*(h,k)\exp\{-2\pi inh/k\} \int \sum_{\nu=0}^{\infty} S(\nu)\exp\{2\pi ih'\nu/k\}$$
$$\cdot \exp\{-(\pi/k^2w)(2\nu - 1/12) + \pi w(2n - 1/12)\}\,d\phi.$$

Splitting the sum over ν into two parts $T(n)$ and $R(n)$ according as $\nu = 0$ or $\nu > 0$ respectively we have $S(n,0) = T(n) + R(n)$. Using Rademacher's procedure [5] and employing Theorem 3 we have, since $2\nu - 1/12 > 0$ if $\nu > 0$,

$$R(n) = O(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\}).$$

For $T(n)$ we have

$$T(n) = 2\pi \sum_{k=1}^{N} C(k,n)P(k,n) + O(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\})$$

where $2 \mid k$, and

$$C(k,n) = \sum_{h \bmod k}{}' \omega^*(h,k)\exp\{-2\pi inh/k\}, \tag{4.1}$$

$$P(k,n) = k^{-1}(24n-1)^{-1/2}I_1\{\pi(24n-1)^{1/2}/6k\}. \tag{4.2}$$

$I_1(x)$ is the Bessel function of order one.

Turning now to $S(n,1)$ we have by Theorem 2,

$$S(n,1)=2^{1/2}\sum_{h,k}{}'\chi^*(h,k)\exp\{-2\pi inh/k\}\int\sum_{\nu=0}^{\infty}q(\nu)\exp\{2\pi iH'\nu/k\}$$
$$\cdot\exp\{-(\pi/2k^2w)(2\nu+1/12)+\pi w(2n-1/12)\}d\phi.$$

Since $2\nu+1/12>0$ we have, again utilizing Rademacher's method and Theorem 3,

$$S(n,1)=O(n^{1/3}N^{-1/3+\epsilon}\exp\{2\pi nN^{-2}\}).$$

Letting $N\to\infty$ we have, since $S(n)=T(n)+R(n)+S(n,1)$,

THEOREM 4. *$S(n)=2\pi\sum_k C(k,n)P(k,n)$ where $2\mid k$. $C(k,n)$ and $P(k,n)$ are given by* (4.1), (4.2) *respectively.*

Since $Q(n)=(-1)^nS(n)$ we have as an immediate

COROLLARY. *The number of partitions, $Q(n)$, of a positive integer n into odd and unequal parts is given by the convergent series*

(4.3) $$Q(n)=2\pi\sum_k(-1)^nC(k,n)P(k,n)$$

where $2\mid k$. C and P are given by (4.1), (4.2) *respectively.*

5. The two series for $Q(n)$. It is of interest to compare the series for $Q(n)$ given in the preceding section with that derived in [3]. With this in mind we state the main result of [3].

THEOREM 5.

(5.1) $$Q(n)=2\pi\sum_k A(k,n)L(k,n)+\pi\sum_k B(k,n)M(k,n).$$

In the first sum $4\mid k$, in the second sum $2\nmid k$. A, B, L, M are defined in (5.2)-(5.5).

(5.2) $$A(k,n)=\sum_{h \bmod k}{}'\omega(h,k)\omega^*(h,k^*)\exp\{-2\pi inh/k\},\ k=2k^*;$$

(5.3) $$B(k,n)=\sum_{h \bmod k}{}'\chi(h,k)\chi^*(2h,k)\exp\{-2\pi inh/k\};$$

(5.4) $$L(k,n)=k^{-1}(24n-1)^{-1/2}I_1\{\pi(24n-1)^{1/2}/6k\};$$

(5.5) $$M(k,n)=k^{-1}(24n-1)^{-1/2}I_1\{\pi(24n-1)^{1/2}/12k\}.$$

We see immediately that (4.3) has several computational advantages over (5.1). In the first place, (4.3) involves only one infinite series with

the same rule of formation applying to each term. In (5.1) two infinite series are involved with (apparently) different rules of formation for their terms. Also, the exponential sums appearing in (4.3), involving the reciprocal of $\omega(h,k)$ are much simpler in form than those of (5.1) which in each term of $A(k,n)$ and $B(k,n)$ involve the product of exponentials and also require the introduction of k^* and $2h$.

We shall conclude by showing explicitly the term by term equality of (4.3) and (5.1). The proof is not quite trivial. We first note that if $4 \mid k$ then obviously $P(k,n) = L(k,n)$, while if h runs through a reduced residue system modulo k then so does $h - k^*$. Since $(-1)^n \exp\{-2\pi inh/k\} = \exp\{-2\pi in(h-k^*)/k\}$ we see from (4.1) and (5.2) that $(-1)^n C(n,k) = A(k,n)$ if $\omega^*(h,k) = \omega(h-k^*,k)\omega^*(h-k^*,k^*)$. From (2.1) this last equality is equivalent to

$$\exp\{-\pi i\Sigma((\mu/k))((h\mu/k))\} \\ = \exp\{\pi i\Sigma((\mu/k))((h\mu/k-\mu/2))\}\exp\{-\pi i\Sigma((\nu/k^*))((h\nu/k^*))\}$$

where $\mu = 1,3,\cdots,k-1$ and $\nu = 1,3,\cdots,k^*-1$. From the properties of $((y))$ we easily verify that $((h\mu/k-\mu/2)) = ((h\mu/k+1/2))$ and $\Sigma((\nu/k^*))((h\nu/k^*)) = \Sigma((\mu/k))((h\mu/k^*))$. Therefore, it follows that we must prove that

$$\exp\{\pi i\Sigma((\mu/k))\{((h\mu/k+1/2)) + ((h\mu/k)) - ((h\mu/k^*))\}\} = 1,$$

or $\exp\{\pi i\Sigma((\mu/k))([2h\mu/k] - [h\mu/k] - [h\mu/k+1/2])\} = 1$. This last equality holds since if the fractional part of y is less than $1/2$ we have $[y+1/2] = [y]$ and $[2y] = 2[y]$, while if the fractional part of y is greater than or equal to $1/2$ we have $[y+1/2] = [y]+1$ and $[2y] = 2[y]+1$.

If $k \equiv 2 \pmod 4$ then $P(k,n) = M(k^*, n/2)$, while if h runs through a reduced residue system modulo k then $(h-k^*)/2$ runs through a reduced residue system modulo k^*. Since

$$(-1)^n \exp\{-2\pi inh/k\} = \exp\{-2\pi in(h-k^*)/2k^*\}$$

it follows from (4.1) and (5.3) that $(-1)^n C(n,k) = B(k^*,n)$ if

$$\omega^*(h,k) = \chi((h-k^*)/2,k^*)\chi^*(h-k^*,k^*).$$

From (2.1) and (2.2) the last equality is equivalent to

$$\exp\{-\pi i\Sigma((\mu/k))((h\mu/k))\} \\ = \exp\{\pi i\Sigma((\mu/k))((h\mu/k-\mu/2))\}\exp\{-\pi i\Sigma((\mu/k))((h\mu/k^*))\}$$

where $\mu = 1,3,\cdots,k-1$. This is proved as in the case $4 \mid k$ just discussed.

TEMPLE UNIVERSITY.

REFERENCES.

[1] P. Hagis, Jr., "A problem on partitions with a prime modulus $p \geq 3$," *Transactions of the American Mathematical Society*, vol. 102 (1962), pp. 30-62.

[2] ———, "Partitions into odd summands," *American Journal of Mathematics*, vol. 85 (1963), pp. 213-222.

[3] ———, "Partitions into odd and unequal parts," *American Journal of Mathematics*, vol. 86 (1964), pp. 317-324.

[4] J. Lehner, "A partition function connected with the modulus five," *Duke Mathematical Journal*, vol. 8 (1941), pp. 631-655.

[5] H. Rademacher, "The Fourier coefficients of the modular invariant $J(\tau)$," *American Journal of Mathematics*, vol. 60 (1938), pp. 501-512.

FREE DERIVATION MODULES ON ALGEBRAIC VARIETIES.

By Joseph Lipman.[1,2]

Introduction. The Jacobian criterion for simple points may be formulated in the following way [**5**; § 3]:

Let P be a point on an algebraic variety V/k over a ground field k, which we assume, for simplicity, to be perfect. Let R be the local ring of P on V, and let $D = D_k(R)$ be the module of k-differentials of R. Then in order that P be a simple point of V, it is necessary and sufficient that D be a free R-module.

$D^* = \mathrm{Hom}_R(D, R)$, the dual module of D, may be identified with the module of k-derivations of R into itself. If P is simple on V, so that D is free, then of course D^* is free. It is tempting to ask for the converse: *If D^* is free, is P simple?*

The answer is in the negative when k has characteristic $p \neq 0$, even under the additional assumption that P is normal, a counterexample being given by the origin on the surface $Z^p = XY$ (cf. § 7).

In characteristic zero, however, the question remains open, even when V is a surface in 3-space. By way of encouragement we have an affirmative answer in some special situations, for example when P is the vertex of a cone, or when P is the origin on a surface whose equation is of the form $Z^n = f(X, Y)$ with $f(0,0) = 0$. (cf. § 7).

Our purpose will be to study, for its own sake, the condition that D^* be free. Although we cannot answer the above question, we can still develop some results which may prove useful toward that end. Assuming that D^* is free, we show in § 3 that when k has characteristic zero, P is a normal point. (Thus if V is a curve, P is simple). In § 5 we give some upper bounds on the codimension of the singular locus in the neighborhood of P. We give a technical criterion for determining whether D^* is free in § 6, and apply this criterion in § 7 to a number of specific examples. In an appendix (§ 8)

Received November 30, 1964.

[1] This paper forms part of the author's dissertation, submitted at Harvard University.

[2] The writing of this paper was partially supported by the Air Force Office of Scientific Research Contract AF-AFOSR-442-63, and also, during the summer of 1963, by the Summer Research Institute of the Canadian Mathematical Congress.

we give a simple characterization, in terms of the local codimension of the singular locus, of those points on a complete intersection whose module of differentials is torsion free or reflexive. Some of the techniques used in the proofs have independent interest (cf. § 4, § 6).

In view of the fact that the module of differentials of R has been defined on certain occasions to be D^{**}, the dual of the module of derivations, one might also ask under what conditions D^{**} is free. However, as long as R is a *reduced* ring, it can be seen that $D^{*} \cong D^{***}$, from which it follows that D^{**} is free if and only if D^{*} is free.

We make the convention now that the word "ring" shall mean "non-null commutative ring with identity," and that all modules shall be unitary. We use the terms "finitely generated" and "of finite type" interchangeably. The phrase "V/k is an affine variety" shall mean "k is a field, and $V = \operatorname{Spec} S$ where S is a reduced k-algebra of finite type."

The author wishes to express his gratitude to Professor Zariski, who originally suggested the above question, for a number of conversations out of which many of the ideas in this paper developed, as well as for constant encouragement.

1. **Generalities.** Let V/k be an affine variety over a *perfect* ground field k; thus $V = \operatorname{Spec} S$, where S is a finitely generated k-algebra without nonzero nilpotent elements.

For any k-algebra A, let $D(A)$ be the module of k-differentials of A and let $D^{*}(A) = \operatorname{Hom}_A(D(A), A)$ be the module of k-derivations of A into itself. (For the definition and properties of differential modules see [**5**; § 1]). If $\mathfrak{p}$ is a prime ideal in S, then $D(S_{\mathfrak{p}})$ may be identified with the localization $[D(S)]_{\mathfrak{p}} = D(S) \otimes_S S_{\mathfrak{p}}$. We recall that if A is any ring, if M and N are two A-modules, and if B is a *flat* A-algebra, then the canonical homomorphism

$$\operatorname{Hom}_A(M, N) \otimes_A B \to \operatorname{Hom}_B(M \otimes_A B, N \otimes_A B)$$

is a monomorphism if M is of finite type, and an isomorphism if M is of finite presentation [**3**; p. 39]. $D(S)$ is of finite presentation and $S_{\mathfrak{p}}$ is flat over S; we may therefore identify $[D^{*}(S)]_{\mathfrak{p}}$ with $D^{*}(S_{\mathfrak{p}})$.

If P is a point of V, and $\mathfrak{p}$ is the corresponding prime ideal in S, then we say "D^{*} is free at P" when $D^{*}(S_{\mathfrak{p}})$ is a free $S_{\mathfrak{p}}$-module.

PROPOSITION 1.1. *The points of V at which D^{*} is free form a dense open subset of V.*

Proof. The set of points where D^{*} is free contains the generic point

of any irreducible component of V, since $S_{\mathfrak{q}}$ is a field when $\mathfrak{q}$ is a minimal prime ideal of S. Thus the proposition is a consequence of the general fact that for any ring A, and any A-module N of finite presentation, the prime ideals $\mathfrak{p}$ such that $N_{\mathfrak{p}}$ is a free $A_{\mathfrak{p}}$-module form an open subset of $\operatorname{Spec} A$ (cf. [**3**; p. 137]). q. e. d.

From now on, we will be concerned only with the purely local properties of P which follow from the assumption that D^* is free at P.

PROPOSITION 1.2. *If D^* is free of rank r at P, then every component of V through P has dimension r.*

Proof. Let R $(=S_{\mathfrak{p}})$ be the local ring of P on V. If $\mathfrak{q}$ is a minimal prime ideal of R then $R_{\mathfrak{q}}$ is the function field of a component of V through P, and all such components are accounted for in this way. Moreover $D^*(R_{\mathfrak{q}}) = D^*(R) \otimes_R R_{\mathfrak{q}}$ is a vector space of dimension r over $R_{\mathfrak{q}}$. Since $R_{\mathfrak{q}}$ is a *separable* extension of k (k is perfect), $R_{\mathfrak{q}}$ has transcendence degree r over k. q. e. d.

Again let R be the local ring of P on V. Whether $D^*(R)$ is free or not depends only on the completion of R:

PROPOSITION 1.3. *Let R' be the completion of R. Then $D^*(R')$ is the completion of $D^*(R)$, and consequently $D^*(R)$ is a free R-module if and only if $D^*(R')$ is a free R'-module.*

Proof. Let $\mathfrak{m}$ be the maximal ideal of R'. Let $D' = D(R')/\bigcap_n \mathfrak{m}^n D(R')$. It is known that D' is the completion of $D(R)$, i. e. $D' = D(R) \otimes_R R'$, and that $D^*(R') \cong \operatorname{Hom}_{R'}(D', R')$ [**2**; §§ 2, 3]. (Note: these observations depend on the fact that $D(R)$ is a finitely generated R-module.)

Since R' is a flat R-module, and $D(R)$ is of finite presentation, we have, as above, the isomorphism

$$D^*(R) \otimes_R R' = \operatorname{Hom}_R(D(R), R) \otimes_R R' \cong \operatorname{Hom}_{R'}(D(R) \otimes_R R', R') \cong D^*(R')$$

and the first assertion is proved. Since R and R' are local rings, and R' is a faithfully flat R-algebra, the second assertion is a special case of the following proposition [**3**; p. 53]:

Let A be a ring, and let B be a faithfully flat A-algebra. An A-module F is projective of finite type if and only if the B-module $F \otimes_A B$ is projective of finite type.

(We recall that for modules over local rings, "projective of finite type" means "free of finite type." [**3**; p. 107].)

2. Depth and the conductor; grade and duality. We now give some preliminary results, to be applied in the next section.

Let A be a noetherian ring. A sequence $\{a_1, a_2, \cdots, a_t\}$ of elements in A is a *prime sequence* (of length t) if a_1 is not a zerodivisor in A, and if, for each $i = 2, 3, \cdots, t$, a_i is not a zerodivisor in $A/Aa_1 + Aa_2 + \cdots + Aa_{i-1}$. The *depth* (or *grade*) of a proper ideal $I < A$ is the length of a maximal prime sequence consisting of elements in I, the lengths of any two such sequences being equal [**7**; Thm. 1.3]. For convenience, the ideal A is said to have depth ∞.

Let $\bar{A}$ be the integral closure of A in its total quotient ring. Let $\mathfrak{C}$ be the *conductor* of A in $\bar{A}$, i.e. the annihilator ideal of the A-module $\bar{A}/A$.

PROPOSITION 2.1. *Let I be an ideal in A such that $Aa : I = Aa$ for every nonzerodivisor a in A. Then $\mathfrak{C} : I = \mathfrak{C}$, and consequently if $\mathfrak{C} \neq A$ then every associated prime ideal of $\mathfrak{C}$ has depth $\leqq 1$.*

Proof. Let $c \in A$ be such that $cI \subseteq \mathfrak{C}$. If $x \in \bar{A}$, then $x = b/a$ ($b, a \in A$, and a is a nonzerodivisor) and $cI \cdot b/a \subseteq A$, i.e. $cb \cdot I \subseteq Aa$. Hence

$$cb \in Aa : I = Aa$$

and $c \cdot b/a \in A$. Thus $c\bar{A} \subseteq A$, i.e. $c \in \mathfrak{C}$ and the first assertion is proved.

If I is an associated prime ideal of $\mathfrak{C}$, then $\mathfrak{C} : I \neq \mathfrak{C}$, so that for some nonzerodivisor a, $Aa : I \neq Aa$, i.e. I is contained in an associated prime ideal $\mathfrak{p}$ of Aa. Since depth $\mathfrak{p} \leqq 1$, the same is true of I. q.e.d.

There is an interesting geometric consequence:

COROLLARY. *Let P be a point of an affine variety V/k, and let A be the local ring of P on V. Then P is a normal point of V (i.e. A is an integrally closed domain) if and only if the singular locus of V is of depth $\geqq 2$ locally at P (i.e. every prime ideal $\mathfrak{p}$ in A such that $A_{\mathfrak{p}}$ is not regular has depth $\geqq 2$).*

Proof. If A is an integrally closed domain, then every prime ideal $\mathfrak{p}$ of depth 1 is of height 1, and $A_{\mathfrak{p}}$ is regular for every such prime ideal.

Conversely, if $\mathfrak{p}$ is a prime ideal in A containing $\mathfrak{C}$, then $A_{\mathfrak{p}}$ is not integrally closed in its total quotient ring [**1**; p. 506], and so $A_{\mathfrak{p}}$ is not regular. Therefore, if the singular locus is locally of depth $\geqq 2$, then every prime ideal containing $\mathfrak{C}$ has depth $\geqq 2$, whence, by the proposition, $\mathfrak{C} = A$, so that A is integrally closed in its total quotient ring. Since A is a *reduced* local ring, A is an integral domain. (Otherwise, there is an idempotent e in the

total quotient ring of A, $e \neq 0$, $e \neq 1$, and e is necessarily in A since e satisfies a relation of integral dependence: $e^2 - e = 0$. Since e and $e - 1$ are zero-divisors in A, they are both nonunits in A, which is impossible). q. e. d.

In particular, if A is a Macaulay ring (for example if V is a complete intersection locally at P), then P is normal if and only if no singular subvariety of V of codimension 1 passes through P. For, A being a Macaulay local ring, the depth of a prime ideal $\mathfrak{p}$ in A is the same as the height of $\mathfrak{p}$, i. e. as the codimension of the subvariety of V corresponding to $\mathfrak{p}$.

* * *

Let A be a ring, M an A-module, and let "$*$" denote the functor "dual" so that $M^* = \mathrm{Hom}_A(M, A)$. Let $M^{**} = (M^*)^*$ be the *bidual* of M.

The canonical bilinear map of $M^* \times M$ into A taking the pair (ω, x) into $\omega(x)$ defines a natural homomorphism $f\colon M \to M^{**}$. The "naturality" of f means, explicitly, that for any homomorphism of modules $h\colon M \to N$, the following diagram commutes:

$$\begin{array}{ccc} M & \xrightarrow{f_M} & M^{**} \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h^{**}} \\ N & \xrightarrow{f_N} & N^{**} \end{array}$$

If f is an isomorphism, we say that M is *reflexive*. Any finitely generated free module is reflexive; a direct summand of a reflexive module is reflexive; hence, any finitely generated projective module is reflexive.

The kernel of f consists of those elements of M which are annihilated by every A-homomorphism of M into A. It follows that if M is a submodule of a free module, in particular if M is a projective module, then f is a monomorphism. It also follows that the dual map of $M \to f(M)$ is an isomorphism of $[f(M)]^*$ onto M^*, so that these two modules may be identified.

We denote the cokernel of f by M^{**}/M, even if f is not injective. Thus we have an exact sequence

$$0 \to f(M) \to M^{**} \to M^{**}/M \to 0$$

Applying "$*$" and taking note of the preceding identification we get the exact sequence

$$\begin{aligned} &0 \to (M^{**}/M)^* \to (M^{**})^* \xrightarrow{f^*} M^* \to \\ &\to \mathrm{Ext}^1(M^{**}/M, A) \xrightarrow{h} \mathrm{Ext}^1(M^{**}, A) \end{aligned} \tag{1}$$

which we will use in a moment.

Since $r = \dim T + \text{tr.}\deg_k T/\mathfrak{q}T$, we have $s \leqq \dim T$. If $y_1, y_2, \cdots, y_s$ form part of a regular system of parameters of T then there are derivations $d_1', d_2', \cdots, d_s'$ of T into T such that $(d_i'y_j)$ is the unit $s \times s$ matrix. Extending the d_i' to derivations of $R_\mathfrak{p}$ into $R_\mathfrak{p}$, we complete the proof.

4. Derivations with invertible determinants.

THEOREM 2. *Let A be a ring containing a field of characteristic zero, and let $\mathfrak{m}$ be an ideal in A such that A is a complete Hausdorff space in its $\mathfrak{m}$-adic topology. Suppose there exist derivations $d_1, d_2, \cdots, d_s$ of A into A and elements $x_1, x_2, \cdots, x_s$ in $\mathfrak{m}$ such that the matrix (d_ix_j) is invertible.*

Then there is a subring B of A such that

1) $x_1, x_2, \cdots, x_s$ are analytically independent over B.
2) A is equal to the power series ring $B[[x_1, x_2, \cdots, x_s]]$.
3) B contains the subring A_d of A on which all the d_i vanish (i.e. $A_d = \{z \in A \mid d_iz = 0 \text{ for all } i\}$).

Proof. Let (b_{ij}) be the inverse matrix of (d_ix_j); replacing d_i by $\bar{d}_i = \sum_w b_{iw}d_w$ affects neither the hypotheses nor the conclusions of the theorem, and $\bar{d}_ix_j = \delta_{ij}$; we may therefore assume from the outset that (d_ix_j) is the unit matrix.

As a first step, let B_1 be the subring of A on which d_1 vanishes, and set $d = d_1$, $x = x_1$. Thus $dx = 1$ and $x_2, x_3, \cdots, x_s \in B_1$. For any element y in A set

$$y^\# = y - xdy + (x^2/2!)d^2y - (x^3/3!)d^3y + \cdots.$$

Then $y \to y^\#$ is a ring homomorphism: the identity $(y+z)^\# = y^\# + z^\#$ is obvious, and the identity $(yz)^\# = y^\#z^\#$ is a direct consequence of Leibnitz' rule for repeated differentiation of a product.

If $y^\# = 0$, then $y \in Ax$; moreover $x^\# = 0$; thus the kernel of $\#$ is the principal ideal Ax. We also note that $d(y^\#) = 0$, so that $y^\# \in B_1$, and that $\#$ restricts to the identity on B_1; hence B_1 is the image of $\#$.

Since $(y - y^\#) \in Ax$, and since $x \in \mathfrak{m}$, we see that any element z in A can be written in the form

$$z = z_0 + z_1x + z_2x^2 + \cdots \quad \text{with } z_i \in B_1$$

$$\begin{aligned}(\text{for, } z = z^\# + w_1x = z^\# + (w_1^\# + w_2x)x &= z^\# + w_1^\#x + w_2x^2 \\ &= z^\# + w_1^\#x + w_2^\#x^2 + w_3x^3 \text{ etc. etc.})\end{aligned}$$

Applying d^n, we find that $(d^n z - n!z_n) \in Ax$. Hence

$$0 = (d^n z - n!z_n)^\# = (d^n z)^\# - (n!z_n)^\# = (d^n z)^\# - n!z_n$$

so that $z_n = (1/n!)(d^n z)^\#$ (cf. Taylor's theorem!).

In particular, if $z = 0$, we have $z_0 = z_1 = z_2 = \cdots = 0$, i.e. x is analytically independent over B_1. Thus $A = B_1[[x]]$.

Let $\mathfrak{m}^\#$ be the image of $\mathfrak{m}$ under $\#$. Then $\mathfrak{m}^\# \subseteq \mathfrak{m} + Ax + \mathfrak{m}$. Thus $\mathfrak{m}^\# = \mathfrak{m} \cap B_1$.

Since $B_1 = A/Ax$ contains a field of characteristic zero, and since B_1 is a complete Hausdorff space in its $\mathfrak{m}^\# = \mathfrak{m} \cap B_1$ topology, we can easily complete the proof by induction, after noting that the restrictoin of $(\# \circ d_i)$ to B_1 is a derivation of B_1 into B_1 $(i = 2, 3, \cdots, s)$; that $x_j \in \mathfrak{m} \cap B_1$ $(j = 2, 3, \cdots, s)$; that the matrix $((\# \circ d_i)x_j)$ is the unit $(s-1) \times (s-1)$ matrix; and finally that if $d_i z = 0$ $(i = 1, 2, \cdots, s)$ then z is in B_1 and $(\# \circ d_i)z = 0$ $(i = 2, 3, \cdots, s)$. q. e. d.

Let R be a noetherian local ring with residue field of characteristic zero, let R' be the completion of R, and let k be a subring of R. Suppose there exist k-derivations $d_1, d_2, \cdots, d_s$ of R into R and nonunits $x_1, x_2, \cdots, x_s$ in R such that the matrix $(d_i x_j)$ is invertible. Let

$$\bar{R} = R/Rx_1 + Rx_2 + \cdots + Rx_s,$$

and let $\bar{R}'$ be the completion of $\bar{R}$, so that there is a canonical diagram

$$\begin{array}{ccc} R & \longrightarrow & R' \\ \downarrow & & \downarrow h \\ \bar{R} & \longrightarrow & \bar{R}' \end{array}$$

where the horizontal arrows represent inclusion maps.

COROLLARY 1. *Under the preceding circumstances there is a map $g \colon \bar{R}' \to R'$ such that hg is the identity map of $\bar{R}'$ (hence g is injective), and such that when $\bar{R}'$ is identified with $g(\bar{R}')$, we have*

1) $k \subseteq \bar{R}$.

2) *$x_1, x_2, \cdots, x_s$ are analytically independent over $\bar{R}'$.*

3) *R' is the power series ring $\bar{R}'[[x_1, x_2, \cdots, x_s]]$.*

Proof. The d_i, being uniformly continuous, may be extended to derivations of R' into R'. Thus, by the theorem, $R' = B[[x_1, x_2, \cdots, x_s]]$, with $k \subseteq B$. If π is the projection of R' onto B, then $g = \pi h^{-1}$ is easily seen to

be a well-defined isomorphism of $\bar{R}'$ onto B having the required properties (as a map of $\bar{R}'$ into R'). q. e. d.

R and B being as above, we note that if R has Krull dimension s, and if R is analytically unramified, then $B = R'/R'x_1 + R'x_2 + \cdots + R'x_s$ is a reduced local ring of dimension zero, i. e. a field. Hence R is a regular local ring and $x_1, x_2, \cdots, x_s$ are regular parameters. This remark applies in particular to the ring $R_{\mathfrak{p}}$ of § 3; thus *the proof of Theorem* 1 *is complete.*

Geometrically, Corollary 1 may be interpreted as follows. If R is the local ring of a point P on an affine variety V/k over a field k of characteristic zero, then R is analytically unramified, and therefore, by the corollary, $\bar{R}$ is analytically unramified. Hence $\bar{R}$ is the local ring of P on an affine subvariety $\bar{V}$ of V and the corollary states that some neighborhood of P on V is analytically the direct product of a neighborhood of P on $\bar{V}$ with an open subset of the affine space k^s. In particular, the singularity of P on V is completely determined by that of P on $\bar{V}$. For further developments in this direction we refer the reader to a paper of Zariski on the subject of "equisingularity" [**9**].

With regard to the study of free D^*, Corollary 1 can be supplemented by a *reduction principle*:

PROPOSITION 4.1. *In the situation described by Corollary* 1, *assume further that the R-module $D(R) = D_k(R)$ of k-differentials of R is finitely generated. Then the $\bar{R}$-module $D(\bar{R}) = D_k(\bar{R})$ is finitely generated, and $D^*(\bar{R})$ is free if and only if $D^*(R)$ is free.*

Proof. If π is chosen as in the proof of the corollary, and $\bar{R}$ is identified with $g(\bar{R})$, then the restriction of π to R is the canonical map of R onto $\bar{R}$, and π is the identity on k. Thus $D(\bar{R})$ is a homomorphic image of $D(R)$ ($D(\bar{R})$ being thought of as an R-module via the map $R \rightarrow \bar{R}$), and it follows that $D(\bar{R})$ is a finitely generated $\bar{R}$-module.

The proof of Proposition 1.3 shows then that $D^*(\bar{R})$ is free if and only if $D^*(\bar{R}')$ is free, and similarly that $D^*(R)$ is free if and only if $D^*(R')$ is free. Also $D^*(\bar{R}')$, $D^*(R')$, are modules of finite type (over $\bar{R}'$, R', respectively). Thus the proposition is a consequence of the following lemma:

LEMMA. *Let B be a ring, let k be a subring of B, and let A be the power series ring $B[[X_1, X_2, \cdots, X_s]]$. Let "$D^*$" denote "module of k-derivations." Then $D^*(B)$ is a projective B-module of finite type if and only if $D^*(A)$ is a projective A-module of finite type.*

Proof. It is sufficient to deal with the case $s=1$. We set $X_1=X$, so that $A=B[[X]]$. A k-derivation of A into A is uniquely determined by its restriction to B, and by its value at X, both of which can be assigned arbitrarily. More precisely, we can check that $D^*(A)\cong D^*(B,A)\oplus A$ (the first summand being the module of k-derivations of B into A). Thus we may replace $D^*(A)$ by $D^*(B,A)$ in the statement of the lemma.

The projection of A onto B induces an A-homomorphism of $D^*(B,A)$ into $D^*(B)$, and we see easily that this mapping is a surjection with kernel $XD^*(B,A)$. Thus $D^*(B)$ is isomorphic as a B-module to

$$D^*(B,A)/XD^*(B,A)=D^*(B,A)\otimes_A B.$$

Hence if $D^*(B,A)$ is projective of finite type, then so is $D^*(B)$.

Conversely, if $D^*(B)$ is a projective B-module of finite type, then $D^*[X]=D^*(B)\otimes_B B[X]$ is a projective $B[X]$-module of finite type. In its X-adic topology, $B[X]$ is a Hausdorff space with completion A. Therefore $D^*(B,A)$ is a topological $B[X]$-module, and it is not hard to see that $D^*[X]$, with its X-adic topology, is a dense topological submodule of $D^*(B,A)$ (we identify $D^*(B,A)$ with the additive group of "power-series in X with coefficients in $D^*(B)$", and similarly we identify $D^*[X]$ with the additive group of "polynomials in X with coefficients in $D^*(B)$," and then $D^*[X]$ is a dense $B[X]$-submodule of $D^*(B,A)$ in an obvious way . . .). Thus $D^*(B,A)$ is the completion of $D^*[X]$. Since $D^*[X]$ is a direct summand of a free $B[X]$-module of finite type, and since completion "commutes" with finite direct sums, $D^*(B,A)$ is a direct summand of a free A-module of finite type. q. e. d.

(The preceding situation can be described succinctly: $D^*(B,A)$ is the *complete tensor product* $D^*(B)\hat{\otimes}_B A$, when B and $D^*(B)$ have discrete topologies, and A has the X-adic topology.)

When we are dealing with free D^* at a point P on an affine variety V/k, with k of characteristic 0, the local ring of P being R, with maximal ideal $\mathfrak{m}$, Proposition 4.1 allows us to assume that $d'\mathfrak{m}\subseteq\mathfrak{m}$ for every k-derivation d' of R into R (otherwise we can pass to a subvariety $\bar{V}$ of V through P, with $\dim\bar{V}<\dim V$).

For example, if P is a (closed) point on a surface S, and if $d'x$ is a unit for some nonunit x in R, then we may pass to R/xR, which is the local ring of P on some curve C through P. If $D^*(R)$ is free, then by Proposition 4.1, $D^*(R/xR)$ is free, whence (Theorem 1) P is a normal point of C, so that R/xR is regular. It follows then from Corollary 1 that R is regular; hence P is a simple point of S.

We close this section with a complement (which we do not require elsewhere) to Theorem 2.

Complement. In the notation of Theorem 2, assume that $d_i x_j = \delta_{ij}$. A necessary and sufficient condition for B to equal A_d is that the d_i commute with each other.

Proof. Necessity is clear. Conversely, let $b \in B$. Set

$$b^{\#} = \sum_{(\alpha_1, \alpha_2, \cdots, \alpha_s)} ((-x_1)^{\alpha_1} \cdots (-x_s)^{\alpha_s} / \alpha_1 ! \cdots \alpha_s !) d_1^{\alpha_1} \cdots d_s^{\alpha_s}(b)$$

where $(\alpha_1, \alpha_2, \cdots, \alpha_s)$ runs through all s-tuples of nonnegative integers.

Then, assuming that the d_i commute with each other, we verify that $d_j(b^{\#}) = 0$ $(j = 1, 2, \cdots, s)$. Hence $b^{\#} \in A_d \subseteq B$, so that all terms of $b^{\#}$ vanish except that for which $(\alpha_1, \alpha_2, \cdots, \alpha_s) = (0, 0, \cdots, 0)$; i. e. $b^{\#} = b$. Thus $B \subseteq A_d$. q. e. d.

The condition that the d_i commute with each other is always satisfied in the geometric case (cf. remarks following Corollary 1). Thus, in this case, $\bar{R}'$ may be identified with the subring of R' on which all the d_i vanish.

5. Free D^* and the codimension of the singular locus. In [**4**; p. 212] Buchsbaum and Rim have obtained a generalization of Krull's principal ideal theorem:

Let R be a noetherian ring, and let $g: R^t \to R^r$ with $t \geqq r$ be a homomorphism of R-modules. Then $\dim R_\mathfrak{p} \leqq t - r + 1$ *for all minimal primes $\mathfrak{p}$ in* Supp(cokernel of g).

If R is the local ring of a point P on an affine variety V/k, and if $D^*(R)$ is free of rank r, then $D^{**}(R)$ is free of rank r, and the above result implies that each irreducible component of $\operatorname{Supp}(D^{**}(R)/D(R))$ is of codimension $\leqq t - r + 1$, t being the least number of generators of $D(R)$ (localizing at a minimal prime of R, we see easily that $t \geqq r$). If k has characteristic zero, then, as we have seen (§ 3), $\operatorname{Supp}(D^{**}(R)/D(R))$ is the singular locus of R. Thus if P is singular (which we hope it is not) we get a bound on the codimension of the singular locus in the neighborhood of P.

It is known [**5**; p. 174] that the number t is characterized by the fact that the t-th *Fitting ideal* of $D(R)$ is the unit ideal in R, while the $(t-1)$-th Fitting ideal is not. If V/k is immersed in affine n-space, i. e. if V is defined by an ideal $I = (f_1, f_2, \cdots, f_s)$ in the polynomial ring

$k[X_1, X_2, \cdots, X_n]$, with $I = \sqrt{I}$, then we have an $s \times n$ Jacobian matrix $(\bar{f}_{ij})$, where $\bar{f}_{ij}$ is the I-residue of the partial derivative $\partial f_i/\partial X_j$, and for any integer q the q-th Fitting ideal of $D(R)$ is generated by the images in R of the $(n-q) \times (n-q)$ subdeterminants of the Jacobian matrix (by convention, such a subdeterminant vanishes if $q < n - s$, and is the identity if $q \geqq n$). Thus $(n-t)$ is the *rank of the Jacobian at* P.

If r is the dimension of V at P, we can associate with P the nonnegative integer $\delta_P = (n-r) - (\text{rank of the Jacobian at } P)$. $\delta_P = t - r$ depends only on R, and not on the particular immersion of V. (The definition of δ_P has nothing to do with the assumption that $D^*(R)$ is free. However, if $D^*(R)$ is free of rank r, then r is indeed the dimension of V at P (Proposition 1.2). In summary:

PROPOSITION 5.1. *Let P be a point of an affine variety V/k over a field of characteristic zero. If D^* is free at P, then every irreducible component of the singular locus which passes through P has codimension $\leqq 1 + \delta_P$ on V.*

If V is a hypersurface in affine n-space, so that $r = \dim V = n-1$, then $\delta_P \leqq 1$ and Proposition 5.1 shows that the components of the singular locus have codimension $\leqq 2$ at P. Since P is normal (Theorem 1) each component actually is of codimension 2 at P. A more general result has been proved by S. Lichtenbaum and M. Schlessinger for complete intersections:

PROPOSITION 5.2. *Let P be a point of an affine variety V/k over a perfect ground field k. Assume further that V/k is a complete intersection locally at P. If D^* is free at P then each component of the singular locus passing through P has codimension $\leqq 2$ on V.*

Proof. Let R be the local ring of P on V. Since V is locally a complete intersection, $D(R)$ has homological dimension $\leqq 1$. If $\mathfrak{p}$ is a prime ideal in R, then $R_\mathfrak{p}$ is regular if and only if $D(R_\mathfrak{p})$ is a free $R_\mathfrak{p}$-module; but $D(R_\mathfrak{p})$ also has homological dimension $\leqq 1$; thus $D(R_\mathfrak{p})$ is free if and only if

$$0 = \mathrm{Ext}_{R_\mathfrak{p}}{}^1(D(R_\mathfrak{p}), R_\mathfrak{p}) = [\mathrm{Ext}_R{}^1(D(R), R)]_\mathfrak{p}.$$

In other words the singular locus of R is $\mathrm{Supp}(\mathrm{Ext}^1(D(R), R))$.

Again, $D(R)$ has homological dimension $\leqq 1$; thus there is an exact sequence

$$0 \to F \to G \to D(R) \to 0$$

with F, G, free R-modules of finite type. Applying $*$, we get the exact sequence

$$0 \to D^*(R) \to G^* \to F^* \to \mathrm{Ext}^1(D(R), R) \to 0$$

By assumption, $D^*(R)$ is free; hence $\mathrm{Ext}^1(D(R),R)$ has homological dimension $\leqq 2$.

Now, Theorem 1.1 of [7] implies that if M is a module of finite type over a noetherian ring, and if $\mathfrak{p}$ is an associated prime ideal of M, then depth $\mathfrak{p} \leqq$ homological dimension of M. Applying this result to the minimal prime ideals of $\mathrm{Supp}(\mathrm{Ext}^1(D(R),R))$, we see that each component of the singular locus of R has depth $\leqq 2$. Since R is a Macaulay ring, each such component has codimension $\leqq 2$. q. e. d.

Remark 1. An alternative description of δ_P (cf. Proposition 5.1) is obtained as follows. If $\mathfrak{m}$ is the maximal ideal of R, then it is well-known [2; § 1] that the universal derivation $d\colon R \to D(R)$ gives rise to an exact sequence of vector spaces over $R/\mathfrak{m}$:

$$0 \to \mathfrak{m}/\mathfrak{m}^2 \to D(R)/\mathfrak{m}D(R) \to D(R/\mathfrak{m}) \to 0$$

The dimension of the vector space $D(R)/\mathfrak{m}D(R)$ is the number t, while the dimension of the space $D(R/\mathfrak{m})$ is the k-dimension of the point P. If V/k is assumed to have dimension r at P, then it follows that

$$r = \dim_{R/\mathfrak{m}} D(R/\mathfrak{m}) + \dim R$$

and that

$$t - r = \delta_P = \dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2 - \dim R.$$

In particular, Proposition 5.1 shows that if P is an isolated singular point of V, and if D^* is free at P, then $2 \dim R \leqq 1 + \dim_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2$.

Remark 2. Let $D = D(R)$ and assume that D^* is free of rank r. The canonical map $D \to D^{**}$ gives rise to a canonical map of symmetric algebras $g\colon S(D) \to S(D^{**})$. We note that $S(D^{**})$ is a polynomial ring in r variables over R.

For each n, let S_n denote the n-th homogeneous component of the symmetric algebra, so that g induces $g_n\colon S_n(D) \to S_n(D^{**})$. If P is an *isolated* singular point of V, and if the ground field k has characteristic zero, then $\mathrm{Supp}(D^{**}/D)$ contains only the maximal ideal of R, and it follows that the cokernel of g_n has finite length $L(n)$ for each n. It is shown in [4; § 3] that $L(n)$ is a polynomial in n for large n, the degree of the polynomial being $r - 1 + \dim R$.

Proposition 5.2 implies therefore that if V is a complete intersection locally at P, then the polynomial $L(n)$ has degree $r + 1$ (since P is normal). A more direct proof of this fact might yield information about $L(n)$ in the

general case (when V is not necessarily a complete intersection), and so lead to an improvement of the estimate in Proposition 5.1.

6. Linear equations and free duals. The following proposition on linear equations with solutions in a module will lead to a characterization of free dual modules (Theorem 3) and, in particular, to a useful criterion for free D^* (Proposition 6.2).

PROPOSITION 6.1. *Let A be a ring, let M be an A-module, and let (a_{ij}) be an $n \times r$ matrix with entries in A. Let I be the ideal generated by the $r \times r$ subdeterminants of (a_{ij}) (if $r > n$, an $r \times r$ subdeterminant is defined to be zero). Then the system of homogeneous linear equations*

$$\text{(S)} \qquad \sum_j a_{ij} X_j = 0 \qquad (i = 1, 2, \cdots, n; j = 1, 2, \cdots, r)$$

has a nontrivial solution in M if and only if I annihilates some nonzero element of M.

Proof. We can always reduce the case where $r > n$ to the case where $r \leqq n$: enlarging the matrix by setting $a_{ij} = 0$ for $i = n+1, n+2, \cdots, r$ affects neither I nor the set of solutions of (S). We assume therefore that $r \leqq n$. Then Cramer's rule implies the necessity of the given condition.

For fixed n, we prove sufficiency by induction on r. If $r = 1$, there is nothing to prove. If $r > 1$, let m be a nonzero element in M which is annihilated by all $r \times r$ subdeterminants of (a_{ij}) (i.e. which is annihilated by I). If m is annihilated by all $(r-1) \times (r-1)$ subdeterminants of (a_{ij}), then by the inductive hypothesis, the truncated system $\sum_j a_{ij} X_j = 0$ $(i = 1, 2, \cdots, n; j = 1, 2, \cdots, r-1)$ has a nontrivial solution; hence, *a fortiori,* (S) has a nontrivial solution.

We may therefore assume that $b_1 m \neq 0$, $b_1, b_2, \cdots, b_r$ being the cofactors of some row in some $r \times r$ submatrix of (a_{ij}). But $\sum_j a_{ij} b_j = \pm c_i$, where for each i, c_i either is an $r \times r$ subdeterminant of (a_{ij}), or is zero. Hence

$$\sum_j a_{ij}(b_j m) = \pm c_i m = 0$$

and we have a nontrivial solution of (S). q.e.d.

We mention, without proof, the following corollary (which we will not use elsewhere):

Let A be a ring, let N be an A-module of finite presentation, with annihilator I', and let M be any A-module. Then

a) $\mathrm{Hom}_A(N, M) = 0$ iff $\mathrm{Hom}_A(A/I', M) = 0$.

b) $N \otimes_A M = 0$ iff $A/I' \otimes_A M = 0$.

[(b) follows directly from (a) via the identity

$$\mathrm{Hom}(N \otimes M, M/I'M) = \mathrm{Hom}(N, \mathrm{Hom}(M, M/I'M))].$$

We proceed to the characterization of free duals.

With any $r \times n$ matrix (a_{ij}), the a_{ij} being elements of a ring A, we associate an A-module of finite presentation, viz. the module with generators $e_1, e_2, \cdots, e_n$ subject to the relations $\sum_j a_{ij} e_j = 0$ $(i = 1, 2, \cdots, r)$. If N is any A-module, then we say that N is *torsion free* if every element of A which is a zerodivisor in N is a zerodivisor in A (cf. § 8).

LEMMA. *Let (a_{ij}) be an $r \times n$ matrix with entries in a ring A, and let N be the associated module. Let I be the ideal generated by the $r \times r$ subdeterminants of (a_{ij}) $(I = (0)$ if $r > n)$. Then*

1) *The rows of (a_{ij}) are linearly independent over A if and only if $(0) : I = (0)$ in A, and if this is so, then*

2) *N is torsion free if and only if $(a) : I = (a)$ for any nonzerodivisor a in A.*

If A is noetherian, then these two conditions on I mean precisely that depth $I \geqq 2$.

Proof. To say that the rows of (a_{ij}) are linearly independent is to say that the system of equations $\sum_j a_{ij} X_i = 0$ has no nontrivial solution in A. By Proposition 6.1 this is equivalent to $(0) : I = (0)$.

If the rows are linearly independent, then "N is torsion free" means: if a is a nonzerodivisor in A, and $(b_1, b_2, \cdots, b_n)$ is a vector with entries in A, and if $a(b_1, b_2, \cdots, b_n)$ is a linear combination with coefficients $x_i \in A$ of the rows of (a_{ij}), then $(b_1, b_2, \cdots, b_n)$ is already such a linear combination; i. e. a divides each x_i.

This condition may be restated as: the system of equations $\sum a_{ij} X_i = 0$ has no nontrivial solution in the A-module $A/(a)$. By Proposition 6.1, this means that $(a) : I = (a)$.

If A is noetherian, then $(0) : I = (0)$ implies that I contains a nonzerodivisor, say a, and then $(a) : I = (a)$ shows that I contains a nonzerodivisor modulo (a). Hence depth $I \geqq 2$. Conversely, if depth $I \geqq 2$, then $(a) : I = (a)$ for any nonzerodivisor a: this is obvious if a is a unit in A,

and otherwise any associated prime ideal of (a) has depth 1, and therefore does not contain I. Similarly $(0):I=(0)$.

COROLLARY. *Let M be a free A-module of finite rank r, let $f_1, f_2, \cdots, f_r$ be linearly independent elements of M, and let Q be the submodule of M generated by the f's. Then M/Q is annihilated by a nonzerodivisor in A.*

Proof. The module M/Q is associated with an $r\times r$ matrix with linearly independent rows. If I is the determinant of this matrix, then $(0):I=(0)$ and I annihilates M/Q (Cramer's rule).

(Conversely, it can be shown that if $f_1, f_2, \cdots, f_r$ are *arbitrary* elements of M, and if M/Q is annihilated by a nonzerodivisor, then $(0):I=(0)$, whence $f_1, f_2, \cdots, f_r$ are linearly independent).

Let A be a ring with total quotient ring K, and let A^0 be the subset of A consisting of the zero element along with all the nonzerodivisors in A. Let M be an A-module and let $M^*=\mathrm{Hom}_A(M,A)$. If $f_1, f_2, \cdots, f_r \in M^*$, and $x_1, x_2, \cdots, x_n \in M$, let $I[f_ix_j]$ be the ideal generated by the $r\times r$ subdeterminants of the matrix (f_ix_j).

THEOREM 3. *With the above notation, assume that M is finitely generated, and let $f_1, f_2, \cdots, f_r \in M^*$. The following statements are equivalent:*

1) *M^* is free and $f_1, f_2, \cdots, f_r$ form a free basis.*
2) *$M^*\otimes_A K$ is a free K-module of rank r, and if $x_1, x_2, \cdots, x_n$ are elements in M which generate M, then $(a):I[f_ix_j]=(a)$ for any a in A^0.*
3) *$M^*\otimes_A K$ is a free K-module of rank r, and there exist elements $x_1, x_2, \cdots, x_n$ in M such that $(a):I[f_ix_j]=(a)$ for any a in A^0.*

Proof. 2) obviously implies 3).

Assume that 3) holds. By enlarging the set $\{x_1, x_2, \cdots, x_n\}$ if necessary, we may assume that $x_1, x_2, \cdots, x_n$ generate M. Let A^n be a free A-module with basis $e_1, e_2, \cdots, e_n$, and let g be the map of A^n onto M such that $g(e_i)=x_i$ $(i=1,2,\cdots,n)$. Thus we have an exact sequence

$$F\to A^n\xrightarrow{g} M\to 0$$

where F is a free A-module. Applying $*$, we get the exact sequence

$$0\to M^*\xrightarrow{g^*}(A^n)^*\to F^*$$

so that M^* may be identified with a submodule of $(A^n)^*$. Note that $(A^n)^*/M^*$ is isomorphic to a submodule of F^*; since F^* is a direct product of copies of A, F^* is torsion free, and therefore $(A^n)^*/M^*$ is torsion free.

We identify any element ω of $(A^n)^*$ with the vector $(\omega e_1, \omega e_2, \cdots, \omega e_n)$. Then $f_i = g^*(f_i) = (f_i x_1, f_i x_2, \cdots, f_i x_n)$ for $i = 1, 2, \cdots, r$. Thus if Q is the submodule of M^* generated by $f_1, f_2, \cdots, f_r$, then $(A^n)^*/Q$ is the module associated with the matrix $(f_i x_j)$. According to the lemma, then, the condition "$(a) : I[f_i x_j] = (a)$ for any a in A^0" means precisely "$f_1, f_2, \cdots, f_r$ are linearly independent (over A), and $(A^n)^*/Q$ is torsion free."

To establish 1), it will be sufficient therefore to show that $Q = M^*$, i.e. $M^*/Q = 0$. Since $M^*/Q \subseteq (A^n)^*/Q$ which is torsion free, it is even enough to show that each element of M^*/Q is annihilated by a nonzerodivisor of A, i.e. that $(M^*/Q) \otimes_A K = 0$, i.e. that $(M^* \otimes_A K/Q \otimes_A K) = 0$.

It is clear that M^* is torsion free, and it follows that the canonical map $M^* \rightarrow M^* \otimes_A K$ is injective. Hence the images of $f_1, f_2, \cdots, f_r$ in $M^* \otimes_A K$, which generate the submodule $Q \otimes_A K$, are linearly independent over K. Since $M^* \otimes_A K$ is, by assumption, a free K-module of rank r, and since every nonzerodivisor in K is a unit in K, the corollary of the lemma shows that indeed $(M^* \otimes_A K/Q \otimes_A K) = 0$. Thus 3) implies 1).

Forgetting 3), assume now that 1) holds, and apply the preceding considerations to any set of elements $x_1, x_2, \cdots, x_n$ which generate M. In this case, $Q = M^*$, and, having remarked that $(A^n)^*/M^*$ is torsion free, we see that 2) holds by referring to the equivalent conditions set in quotation marks three paragraphs back. This completes the proof.

As an immediate corollary of Theorem 3 and the last assertion of the lemma, we have:

Proposition 6.2. *Let R be the local ring of a point P on an affine variety V/k over a perfect ground field k. Assume that V/k is locally, at P, equidimensional of dimension r. Then the k-derivations $d_1, d_2, \cdots, d_r$ of R into R form a free basis of $D^*(R)$ if and only if there exist elements $x_1, x_2, \cdots, x_n$ in R such that the ideal generated by the $r \times r$ subdeterminants of the matrix $(d_i x_j)$ has depth $\geqq 2$.*

Moreover, if $d_1, d_2, \cdots, d_r$ do form a free basis for $D^*(R)$, then the ideal generated by the $r \times r$ subdeterminants of the matrix $(d_i x_j)$ has depth $\geqq 2$ whenever $x_1, x_2, \cdots, x_n$ are such that the k-differentials $dx_1, dx_2, \cdots, dx_n$ generate $D(R)$.

Remark 6.3. Clearly Proposition 6.2 also holds if R is the *coordinate*

ring of the affine variety V/k, provided that all irreducible components of V have dimension r.

7. **Examples.** a) If V is an affine curve over a field of characteristic zero, then, by Theorem 1, D^* is free at a point P of V if and only if P is simple. On the other hand, consider the irreducible plane curve C defined over a perfect ground field k of characteristic $p \neq 0$ by the equation $f(X,Y) = X^p - Y^{p+1} = 0$. We have $f_X = 0$, $f_Y = Y^p$. Thus if $R = k[x,y]$ is the local ring of the origin, then there is a derivation $\bar{d}$ of R into R with $\bar{d}x = 1$, $\bar{d}y = 0$. It follows from Remark 6.3 (or it may be checked directly) that $D^*(R)$ is free with generator $\bar{d}$. Thus D^* is free at every point of C. Nevertheless, the origin is not a normal point of C.

An example where D^* is not globally free as above is provided by the curve $Y^2 + X^p + X^{p+1} = 0$ over a perfect field of characteristic $p > 2$. If R is the local ring of the origin, then there os a derivation $\bar{d}$ of R into R with $\bar{d}x = 2 + 2x$, $\bar{d}y = y$. By Proposition 6.2 (or directly) $D^*(R)$ is a free module with generator $\bar{d}$, whereas the origin is a singular point. (The fact that D^* is not globally free is seen by remark 6.3 and by consideration of the points $(0,0)$, $(-1,0)$ on the curve).

b) Consider the surface defined over a perfect ground field k of characteristic $p \neq 0$ by the equation $f(X,Y,Z) = XY - Z^p = 0$. We have $f_X = Y$, $f_Y = X$, $f_Z = 0$. The origin is the only singular point, and by the corollary to Proposition 2.1, the origin is a normal point. Thus, the coordinate ring $R = k[x,y,z]$ is integrally closed, and the ideal (x,y) in R has depth 2. By Remark 6.3, the two derivations d_1, d_2 such that: $d_1x = 0, d_1y = 0, d_1z = 1$; $d_2x = -x$, $d_2y = y$, $d_2z = 0$; form a free basis of $D^*(R)$. So we can have free D^* in the presence of singular points, even under the assumption of normality.

From now on, we restrict ourselves to a fixed ground field k of *characteristic zero.*

c) Let P be the origin on a cone K in affine 3-space (over k), K being given by $f(X,Y,Z) = 0$ where f is a homogeneous polynomial without multiple factors. Zariski has shown (unpublished notes) that D^* is free at P only if P is a simple point of K (i.e. only if K is a *plane*). The proof is given here with his permission.

Assume that P is a singular point of K. If D^* is free at P, P is normal (Theorem 1), and it follows that P is the only singular point of K (otherwise the line joining P to a singular point would be a multiple curve through

P). If f_X, f_Y, f_Z all vanish at some point P' in the affine space then, by Euler's theorem on homogeneous polynomials, f also vanishes at P', and P' is a singular point of K; hence $P'=P$. In other words, (X,Y,Z) is the only associated prime ideal (in $k[X,Y,Z]$) of the ideal (f_X,f_Y,f_Z). Hence no associated prime ideal of (f_Z,f_Y) is a minimal prime ideal, and it follows that $(f_Z):(f_Y)=(f_Z)$. It also follows, by Macaulay's theorem, that (f_Z,f_Y) is pure one-dimensional so that $(f_Z,f_Y):(f_X)=(f_Z,f_Y)$. We shall make use of these observations below.

Let R be the local ring of P on K, and let x, y, z be the traces of X, Y, Z on K, so that $k[x,y,z]$ is the coordinate ring of K. $k[x,y,z]$ is a subring of R since every irreducible component of a cone contains the vertex. We identify any derivation $\bar{d}$ of R into R with the vector $(\bar{d}x,\bar{d}y,\bar{d}z)$. Thus the derivations of R into R are the vectors (a,b,c) with a, b, c in R such that

$$af_x+bf_y+cf_z=0 \qquad\qquad f_x=f_X(x,y,z) \text{ etc.}$$

The derivations (a,b,c) with a, b, c in $k[x,y,z]$ span the R-module of all derivations of R into R. We seek, therefore, the polynomial solutions (A,B,C) of

$$Af_X+Bf_Y+Cf_Z\equiv 0 \pmod{f}$$

These solutions form a $k[X,Y,Z]$-module M which is spanned by the *homogeneous* solutions (since f is homogeneous). So let A, B, C, E be homogeneous polynomials such that

$$Af_X+Bf_Y+Cf_Z=Ef$$

Setting $A_1=A-XE/n$, $B_1=B-YE/n$, $C_1=C-ZE/n$, where n $=$ degree of f, we get

$$A_1f_X+B_1f_Y+C_1f_Z=0 \tag{2}$$

Hence (X,Y,Z) and the homogeneous solutions of (2) span M. Since $(f_Z,f_Y):(f_X)=(f_Z,f_Y)$, (2) implies that $A_1=Gf_Y+Hf_Z$ (G, H homogeneous polynomials). Subtracting from (A_1,B_1,C_1) the vector $G(f_Y,-f_X,0)$ $+H(f_Z,0,-f_X)$ (which is in M) we may assume $A_1=0$. But then, since $(f_Z):(f_Y)=(f_Z)$, (2) implies that $(0,B_1,C_1)=L(0,f_Z,-f_Y)$ (L a polynomial).

Hence M is spanned by the vectors

$$(X,Y,Z),\ (f_Y,-f_X,0),\ (-f_Z,0,f_X),\ (0,f_Z,-f_Y)$$

and therefore $D^*(R)$ is spanned by the vectors

$$(x,y,z),\ (f_y,-f_x,0),\ (-f_z,0,f_x),\ (0,f_z,-f_y).$$

Thus there is a free basis of $D^*(R)$ among these four vectors. Applying Proposition 6.2 to all the possible pairs among these vectors, we conclude that one of f_x, f_y, f_z is a unit in R; this is impossible since P is not simple.

d) We will treat in some detail the following situation: P is the origin on a surface V defined over the field k (of characteristic zero) by an equation of the form $Z^n - f(X,Y)$ where $f(0,0) = 0$, and $n > 1$. If D^* is free at P, then P is normal (Theorem 1); by the corollary of Proposition 2.1, P is normal if and only if, writing $f(X,Y) = g(X,Y)h(X,Y)$, where $h(X,Y)$ is the product of all those factors of $f(X,Y)$ which do not vanish at $(0,0)$, we find that $g(X,Y)$ has no multiple factors.

We will show that if D^* is free at P, then D^* is free at the origin on the curve defined by $g(X,Y) = 0$. By Theorem 1, the origin is a simple point of this curve, whence P is a simple point of V.

(i) Let $k[x,y,z]$ be the coordinate ring of V, and let S be the local ring of P on V. Let R be the local ring of the origin on the (x,y) plane, i.e. let $R = k[x,y]_{(x,y)}$. Then $R[z] \subseteq S$. On the other hand, we can check that $R[z]$ is a local ring, cf. [8; p. 318], and it follows easily that $R[z] = S$.

(ii) We study the derivations of $S = R[z]$ into itself. Any such derivation is uniquely determined by its restriction to R; thus our problem is to study the derivations of R into $R[z]$ which can be extended to derivations of $R[z]$ into $R[z]$.

Let $\bar{d}$ be a derivation of R into $R[z]$. Then clearly

$$\bar{d} = d_0 + zd_1 + z^2d_2 + \cdots + z^{n-1}d_{n-1}$$

where $d_0, d_1, \cdots, d_{n-1}$ are uniquely determined derivations of R into R. We claim that *$\bar{d}$ can be extended to a derivation of $R[z]$ into $R[z]$ if and only if $d_0, d_1, \cdots, d_{n-2}$ can be extended.* (Note that the derivation $z^{n-1}d_{n-1}$ can always be extended.)

If this is so, then, denoting extensions by upper "e," we will have

$$\bar{d}^e = (d_0)^e + z(d_1)^e + z^2(d_2)^e + \cdots + (z^{n-1}d_{n-1})^e$$

since the derivation on the right is obviously an extension of $\bar{d}$, and since $\bar{d}$ has at most one extension. It follows that *$D^*(S)$ is generated by derivations of the form $(\bar{d}_0)^e$, or $(z^{n-1}\bar{d}_0)^e$,* where $\bar{d}_0$ is a derivation of R into R.

The proof of the claim is straightforward. $\bar{d}$ can be extended if and only if there is an s in S such that $nz^{n-1}s = \bar{d}f$, or equivalently,

$$nfs = z(\bar{d}f) = z(d_0f) + z^2(d_1f) + \cdots + z^{n-1}(d_{n-2}f) + f(d_{n-1}f).$$

Thus s exists if and only if f divides $d_0f, d_1f, \cdots, d_{n-2}f$.

Setting $d_1 = d_2 = \cdots = d_{n-1} = 0$, we see that d_0 can be extended if and only if f divides d_0f; similarly d_i can be extended if and only if f divides d_if $(i = 1, 2, \cdots, n-2)$. Thus, the preceding statement becomes "s exists if and only if $d_0, d_1, \cdots, d_{n-2}$ can be extended." q.e.d.

(iii) Assume now that $D^*(S)$ is a free S-module. Since S is a local ring, any set of generators of $D^*(S)$ contains a free basis $\{d', d''\}$ of $D^*(S)$. By the above results, we may assume that d' is either of the form

$$(\#) \qquad (d_0')^o\colon d_0' \in D^*(R) \text{ and } f \text{ divides } d_0'f$$

or of the form

$$(\#\#) \qquad (z^{n-1}d_0')^o\colon d_0' \in D^*(R)$$

and similarly for d''.

We identify d' with the vector $(d'x, d'y, d'z)$ and d'' with the vector $(d''x, d''y, d''z)$. Then

$$d' = (\alpha_1, \alpha_2, z(d_0'f)/nf) \qquad \alpha_1, \alpha_2 \in R$$

or

$$d' = (z^{n-1}\alpha_1, z^{n-1}\alpha_2, z^{n-1}\cdot z(d_0'f)/nf) \qquad \alpha_1, \alpha_2 \in R$$

according as d' is of the form $(\#)$ or $(\#\#)$. Similar remarks apply to d'', with "β" in place of "α."

Thus we are led to the "derivation matrix"

$$J = (z^{n-1})^t \cdot (z/nf) \begin{pmatrix} \alpha_1 & \alpha_2 & d_0'f \\ \beta_1 & \beta_2 & d_0''f \end{pmatrix}$$

where t $(= 0, 1, \text{or } 2)$ is the number of derivations among d', d'' having the form $(z^{n-1}\cdot\ \cdot)^o$.

Proposition 6.2 states that the 2×2 subdeterminants of J must generate an ideal of depth $\geqq 2$. If $t = 0$, then f divides both $d_0'f$ and $d_0''f$ so that z divides J. If $t = 2$, then $(z^{n-1})^t\cdot(z/nf) = z^{n-1}/n$ and once again z divides J. Hence $t = 1$, and $(z^{n-1})^t\cdot(z/nf)$ is a unit; moreover, we may assume that d' is of the form $(\#)$, so that f divides $d_0'f$.

Now $d_0'f = f_x\alpha_1 + f_y\alpha_2$ and $d_0''f = f_x\beta_1 + f_y\beta_2$, so that the last column of J is a linear combination of the first two. Thus *Proposition* 6.2 *will be satisfied only if* $\alpha_1\beta_2 - \alpha_2\beta_1$ *is a unit in* S (and therefore in R). Hence either $\alpha_1 = d_0'x$ or $\alpha_2 = d_0'y$ is a unit in R. Moreover, since f divides $d_0'f$, d_0' induces a derivation in $R/(f)$, i.e. in $R/(g)$ which is the local ring of the origin on the plane curve $g(X, Y) = 0$. Since, say, α_1 (mod. g) is a unit

in $R/(g)$, we see, by Proposition 6.2, that $D^*(R/(g))$ is a free module with generator d_0' (mod. g). This is what we set out to prove.

e) The method of d) can be extended to more complicated situations. For example, the origin on a 3-fold in 5-space given by equations of the form

$$U^n = f(W,X,Y,Z) \qquad\qquad f(0,0,0,0)=0$$
$$W^m = g(X,Y,Z) \qquad\qquad g(0,0,0)=0$$

has a free D^* if and only if it is a simple point.

f) The purpose of the examples in this section has been to illustrate the conjecture that P is simple if D^* is free at P. We wish to point out that in attempting to prove this conjecture (assuming that k has characteristic zero), we may assume that k is algebraically closed, and that P is a rational point. For, in the first place, Proposition 1.1 shows that if D^* is free at P, then D^* is free at almost every algebraic specialization (over k) of P, so that we may assume that P is algebraic (over k).

Secondly, let $\bar{k}$ be the algebraic closure of k; then we have the canonical projection $V\times_k\bar{k}\to V$. If $\bar{P}$ is any point of $V\times_k\bar{k}$ lying over P, then the local ring $\bar{R}$ of $\bar{P}$ is the localization at one of the maximal ideals of the semi-local ring $R\otimes_k\bar{k}$, R being the local ring of P on V. $\bar{R}$ is a faithfully flat R-algebra.

It is not hard to see that any k-derivation of R into an $\bar{R}$-module M has a unique extension to a $\bar{k}$-derivation of $\bar{R}$ into M. It follows easily that $\bar{D}(\bar{R}) = D(R)\otimes_R\bar{R}$ where "$\bar{D}$" denotes "module of $\bar{k}$-differentials." Hence (cf. proof of Proposition 1.3) $\bar{D}^*(\bar{R}) = D^*(R)\otimes_R\bar{R}$. Since $\bar{R}$ is faithfully flat over R, $\bar{D}(\bar{R})$ (respectively $\bar{D}^*(\bar{R})$) is free if and only if $D(R)$ (respectively $D^*(R)$) is free (cf. proof of Proposition 1.3). This shows that, for the purposes of the conjecture, we may assume $k=\bar{k}$.

8. Appendix: Torsion free and reflexive differential modules. We will indicate a proof of the following facts:

PROPOSITION 8.1. *Let R be the local ring of a point P on an affine variety V/k over a perfect ground field k. Assume that V is locally, at P, a complete intersection. Let $D(R)$ be the R-module of k-differentials of R. Then*

1) *$D(R)$ is torsion free if and only if V is nonsingular in codimension 1 at P, i.e. if and only if P is normal.*

2) *$D(R)$ is reflexive if and only if V is nonsingular in codimension 2 at P.*

Given a ring A with total quotient ring K, and an A-module M, we call the kernel of the canonical map $i\colon M \to M_{(K)} = M \otimes_A K$ the *torsion submodule* of M. Thus the torsion submodule consists of all elements of M which are annihilated by a nonzerodivisor in A. M is torsion free (cf. § 6) if and only if its torsion submodule is (0). One checks that the torsion submodule of M is contained in the kernel of the canonical map $f\colon M \to M^{**}$. Conversely, if K is semisimple (equivalently: if the ideal (0) is a finite intersection of prime ideals in A) then the "naturality" of f gives a commutative diagram

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[d, "i"'] & M^{**} \arrow[d, "i^{**}"] \\ M_{(K)} \arrow[r, "g"] & M_{(K)}^{**} \end{tikzcd}$$

in which g *is* injective, since $M_{(K)}$ is a projective K-module (cf. § 2). Hence, in this case, $\ker f \subseteq \ker i$, so that the torsion submodule *is* the kernel of f.

The next lemma gives further information about the kernel and cokernel of $f\colon M \to M^{**}$.

LEMMA. *Let A be any ring, and let $F_0 \to F_1 \to M \to 0$ be an exact sequence of A-modules, where F_0 and F_1 are projective and of finite type. Let N be the cokernel of the dual map $F_1^* \to F_0^*$. Then there is an exact sequence*

$$0 \to \operatorname{Ext}^1(N, A) \to M \xrightarrow{f} M^{**} \to \operatorname{Ext}^2(N, A) \to 0$$

Proof. One checks that for any zero-sequence $G_0 \to G_1 \to G_2 \to G_3$ there is an exact sequence

$$\begin{aligned} 0 \to (\text{homology at } G_1) \to (\text{cokernel of } G_0 \to G_1) \to \\ \to (\text{kernel of } G_2 \to G_3) \to (\text{homology at } G_2) \to 0 \end{aligned} \tag{3}$$

We are given the exact sequence

$$0 \to M^* \to F_1^* \to F_0^* \to N \to 0. \tag{4}$$

Building an exact sequence

$$F_3 \to F_2 \to M^* \to 0 \tag{5}$$

with F_3 and F_2 projective modules, combining (4) and (5), and dualizing, we get a zero-sequence

$$0 \to N^* \to F_0^{**} \to F_1^{**} \to F_2^* \to F_3^*.$$

Since F_0 and F_1 are reflexive the cokernel of $F_0^{**} \to F_1^{**}$ can be identified with M; also the kernel of $F_2^* \to F_3^*$ is M^{**}. One checks then that (3) gives rise to the desired sequence. q. e. d.

To prove the proposition, we apply the preceding considerations to the case $A = R$, $M = D(R)$. Since $D(R)$ has homological dimension $\leqq 1$, we may assume that $N = \operatorname{Ext}^1(D(R), R)$. K is now semisimple and it follows that $N \otimes_A K = 0$, whence $N^* = \operatorname{Ext}^0(N, R) = 0$.

By the lemma (and the remarks preceding the lemma), $D(R)$ is torsion free iff $\operatorname{Ext}^1(N, R) = 0$, i. e. iff grade $N \geqq 1$, i. e. iff $\operatorname{Supp} N$ has depth $\geqq 1$ (cf. § 2); similarly $D(R)$ is reflexive iff $\operatorname{Ext}^1(N, R) = \operatorname{Ext}^2(N, R) = 0$, i. e. iff $\operatorname{Supp} N$ has depth $\geqq 2$. However, we have seen, in proving Proposition 5. 2, that $\operatorname{Supp} N$ is the singular locus of R. Also, since R is a Macaulay ring, depth and codimension coincide. Thus, in view of the corollary to Proposition 2. 1, all our assertions are proved.

We remark that Proposition 8. 1, applied to the generic point of a component of the singular locus of V, yields an alternative proof of Proposition 5. 2.

HARVARD UNIVERSITY.

REFERENCES.

[1] S. Abhyankar, "A remark on the nonnormal locus of an analytic space," *Proceedings of the American Mathematical Society*, vol. 15 (1964), pp. 505-508.

[2] R. Berger and E. Kunz, "Uber die Struktur der Differentialmoduln von diskreten Bewertungsringen," *Mathematische Zeitschrift*, vol. 77 (1961), pp. 314-338.

[3] N. Bourbaki, Algèbre Commutative, Ch. 1, 2, Actualités Scientifiques et Industrielles 1290, Hermann, Paris, 1961.

[4] D. A. Buchsbaum and D. S. Rim, "A generalized Koszul complex," *Transactions of the American Mathematical Society*, vol. 111 (1964), pp. 197-224.

[5] E. Kunz, "Die Primidealteiler der Differenten in allgemeinen Ringen," *Journal für die Reine und Angewandte Mathematik*, vol. 204 (1960), pp. 165-182.

[6] M. Nagata, "Remarks on a paper of Zariski on the purity of branch loci," *Proceedings of the National Academy of Sciences, U. S. A.*, vol. 44 (1958), pp. 796-799.

[7] D. Rees, "The grade of an ideal or module," *Proceedings of the Cambridge Phi-*

[8] O. Zariski and P. Samuel, *Commutative Algebra*, Vol. 1, Van Nostrand, Princeton, 1958.

losophical Society, vol. 53 (1957), pp. 28-42.

[9] O. Zariski, "Equivalent singularities of plane algebroid curves," *American Journal of Mathematics*, vol. 87 (1965), pp. 507-536.

GENERALIZED CAYLEY TRANSFORMATIONS OF BOUNDED SYMMETRIC DOMAINS.

By JOSEPH A. WOLF [1] and ADAM KORÁNYI.[2]

1. Introduction. This paper is a continuation of [7]; its main subject is the study of the realizations of Hermitian symmetric spaces as Siegel domains of type III. The general definition of such a domain was given by Pjateckiĭ-Šapiro [9] as follows.

Let $\mathfrak{v}_1$, $\mathfrak{v}_2$ and $\mathfrak{v}_3$ be complex vector spaces. Let $\mathfrak{u}_1$ be a real form of $\mathfrak{v}_1$, $\mathfrak{c}$ an open cone in $\mathfrak{u}_1$, and D a bounded domain in $\mathfrak{v}_3$. Given any $W \in D$ let $\Lambda_W^{(1)}: \mathfrak{v}_2 \times \mathfrak{v}_2 \to \mathfrak{v}_1$ be a bilinear form Hermitian with respect to $\mathfrak{u}_1$, let $\Lambda_W^{(2)}: \mathfrak{v}_2 \times \mathfrak{v}_2 \to \mathfrak{v}_1$ be a complex-symmetric bilinear form, and define $\Lambda_W = \Lambda_W^{(1)} + \Lambda_W^{(2)}$. Then the domain

$$\{E_1 + E_2 + E_3 \in \mathfrak{v}_1 + \mathfrak{v}_2 + \mathfrak{v}_3 : \mathrm{Im.}\, E_1 - \mathrm{Re.}\, \Lambda_{E_3}(E_2, E_2) \in \mathfrak{c}\}$$

is called a Siegel domain of type III.

Pjateckiĭ-Šapiro [9] gave a case by case determination of the realizations of the classical irreducible Hermitian symmetric spaces as Siegel domains of type III. In this paper we will determine those realizations for all Hermitian symmetric spaces by a method which is independent of the classification theory. This is closely related to the study of the boundary structure of bounded symmetric domains. In the classical cases that study is due to Pjateckiĭ-Šapiro [9]; in the general case most of the relevant results have been proved by C. C. Moore [8], who combined our partial Cayley transform with the general theory of boundaries due to Furstenberg [3] and Satake [10]. In this paper (Section 4) we give an explicit direct description of the boundary structure. The greater simplicity of our methods, and the fact that many intermediary results from Section 4 are needed in subsequent discussions, are the reasons why those results are included in this paper.

Our method is an extension of the technique of [7]. Making use of the embedding theorems of Borel and Harish-Chandra, we define partial Cayley

Received July 22, 1964.
Revised September 23, 1964.

[1] Partially supported by N. S. F. Grant GP-812.
[2] Partially supported by N. S. F. Grant G-24943.

transformations which carry the bounded domain realization of Harish-Chandra to the various Siegel domains of type III.

In the case of a polycylinder $U^n \subset C^n$ (where U denotes the unit disc in C) a partial Cayley transformation is simply the usual Cayley transformation on some of the factors and the identity transformation on the remaining factors. In the case of a general bounded symmetric domain D in Harish-Chandra realization, it follows from results of Harish-Chandra and is explicitly pointed out by Hermann [5] that D contains a totally geodesic polycylinder U^n with $K \cdot U^n = D$; here n is the rank of D as a symmetric space and K is the isotropy subgroup at the origin of the connected group G^0 of holomorphic automorphisms of D. A partial Cayley transformation of D can be viewed as a natural extension to D of a partial Cayley transformation of U^n.

Sections 2, 3 and 4 contain a considerable amount of expository material, which is included so that the paper can be used by beginners in the subject. In Section 2 we introduce our notation and some definitions. In Section 3 we collect some facts on parabolic subgroups of real Lie groups; these are due to A. Borel and J. Tits [2] and to a conversation between J. Tits and the first-named author. In Section 4 we give an explicit description of the boundary components of D (Theorem 4.8) and compute them in the irreducible cases (Theorem 4.13). We do not reprove [8, Theorems 1 and 2] because Moore's proof is independent of the general theory of boundaries of symmetric spaces.

In Sections 5 and 6 we show that the set of all analytically equivalent ("same type") boundary components is, for each type, both a homogeneous space of K and of G^0; we study the Riemannian geometry and topology of these spaces in some detail. The isotropy subgroup B^Γ of G^0 is transitive on D; this fact is basic in Section 7 where we construct the image of D under the partial Cayley transformation; this Cayley transform is an orbit of a certain conjugate of B^Γ, which we determine explicitly. The resulting domain is a Siegel domain of type III, and in the classical irreducible cases our results specialize to those of Pjateckiĭ-Šapiro.

The results of [7] are degenerate special cases of theorems in the present Sections 5, 6 and 7, but some of our proofs here depend on the results of [7].

2. Notations. As in [7], M will be a Hermitian symmetric space of non-compact type, G^0 its connected group of isometries, and K the isotropy group. G^0 is globally a product of simple Lie groups and M a product of non-compact irreducible Hermitian symmetric spaces. The Lie algebras of G^0 and K are $\mathfrak{g}^0$ and $\mathfrak{k}$, $\mathfrak{g}^C$ is the complexification of $\mathfrak{g}^0$, G^C the adjoint group

of $\mathfrak{g}^C$. G^0 is contained in G^C as the analytic group corresponding to $\mathfrak{g}^0$. The symmetry of $\mathfrak{g}^0$ is denoted by σ; under σ we have the splitting $\mathfrak{g}^0 = \mathfrak{k} + \mathfrak{p}^0$. Let $\mathfrak{p} = i\mathfrak{p}^0$, $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$, G the corresponding analytic subgroup of G^C.

$\mathfrak{h}$ is a Cartan subalgebra in $\mathfrak{k}$; then $\mathfrak{h}^C$ is a Cartan subalgebra in $\mathfrak{g}^C$. The roots of $\mathfrak{g}^C$ which are also roots of k^C are called compact roots. Given a system of simple roots, if $\mathfrak{g}^C$ is simple, there is a unique non-compact simple root. To each root α we have the standard basis elements H_α, E_α. $\mathfrak{p}^+$ and $\mathfrak{p}^-$ are the abelian subalgebras of $\mathfrak{g}^C$ spanned by the positive (resp. negative) non-compact root vectors E_α; P^+ and P^- are the corresponding analytic groups in G^C.

K^C denoting the analytic subgroup corresponding to $\mathfrak{k}^C$, $K^C \cdot P^+$ is a semidirect product. G/K is identified with $G^C/K^C \cdot P^+$ by the identity map of G into G^C; this space is the compact dual of M, and is denoted by M^*. x denotes the identity coset in $M^* = G^C/K^C \cdot P^+$. The orbit $G^0(x)$ is the image of the holomorphic embedding $gK \to g(x)$ of M into M^*. The map $\xi\colon \mathfrak{p}^- \to M^*$, defined by $\xi(E) = \exp(E) \cdot (x)$ is a holomorphic homeomorphism onto a dense open subset; ξ is $\mathrm{ad}(K^C)$-equivariant. $D = \xi^{-1}(G^0(x))$ is a bounded symmetric domain in $\mathfrak{p}^-$; this is the Harish-Chandra realization of M.

The center $\mathfrak{z}$ of $\mathfrak{k}$ contains an element Z such that $\mathrm{ad}(Z)E = \pm iE$ for $E \in \mathfrak{p}^\mp$. $J = \mathrm{ad}(Z)$ is a complex structure on $\mathfrak{p}^0$. A basis of $\mathfrak{p}^0$ is given by the elements

$$X_\alpha{}^0 = E_\alpha + E_{-\alpha}$$
$$Y_\alpha{}^0 = -i(E_\alpha - E_{-\alpha}),$$

where α is non-compact positive. For such α we have the relations

$$JX_\alpha{}^0 = [Z, X_\alpha{}^0] = Y_\alpha{}^0$$
$$JY_\alpha{}^0 = [Z, Y_\alpha{}^0] = -X_\alpha{}^0$$
$$[X_\alpha{}^0, Y_\alpha{}^0] = 2iH_\alpha.$$

We define the elements $X_\alpha = iX_\alpha{}^0$, $Y_\alpha = iY_\alpha{}^0$; these form a basis of $\mathfrak{p}$.

Two roots α and β of $\mathfrak{g}^C$ are called strongly orthogonal if $\alpha \pm \beta$ are not roots. There exists a set Δ of strongly orthogonal positive non-compact roots such that the real span $\mathfrak{a}^0$ of the $X_\alpha{}^0$ $(\alpha \in \Delta)$ is a maximal abelian subalgebra contained in $\mathfrak{p}^0$. $\Delta = \{\delta_1, \cdots, \delta_r\}$ is constructed in [4] as follows: For each j, δ_{j+1} is the lowest positive non-compact root that is strongly orthogonal to $\delta_1, \cdots, \delta_j$. Thus, if $\mathfrak{g}$ is simple, δ_1 is the non-compact simple root. In our proofs we shall calculate with a set Δ constructed in this way. Our results, however, do not depend on the construction of Δ.

For each $\alpha \in \Delta$ we define the 3-dimensional simple subalgebras $\mathfrak{g}_\alpha$, spanned by $\{iH_\alpha, X_\alpha, Y_\alpha\}$, and $\mathfrak{g}_\alpha{}^0$, spanned by $\{iH_\alpha, X_\alpha{}^0, Y_\alpha{}^0\}$. The corresponding analytic subgroups of G^C are G_α and $G_\alpha{}^0$. We define $\mathfrak{h}^- = [\mathfrak{a}^0, J\mathfrak{a}^0]$; then $\mathfrak{h}^- \subset \mathfrak{h}$. $\mathfrak{h}^+$ denotes the orthogonal complement of $\mathfrak{h}^-$ in $\mathfrak{h}$ with respect to the Killing form; $\mathfrak{h}^+$ is the centralizer of $\mathfrak{a}^0$ in $\mathfrak{h}$, and $\mathfrak{t} = \mathfrak{h}^+ + \mathfrak{a}^0$ is a Cartan subalgebra of $\mathfrak{g}^0$, by [7, Proposition 3.8]. We have $Z = Z^0 + Z'$, where $Z^0 = -\frac{1}{2}\sum_{\alpha\in\Delta} H_\alpha \in \mathfrak{h}^-$ and $Z' \in \mathfrak{h}^+$.

For every $\alpha \in \Delta$ we have $c_\alpha = \exp(\pi/4)X_\alpha \in G$; $c = \prod_{\alpha\in\Delta} c_\alpha$ is the *Cayley transform* of M. $\operatorname{ad}(c)$ has order 8 or 4. If it has order 4, M is said to be of *tube type*. This is equivalent to the fact that M can be realized as a tube domain over a self-dual cone (Remark 1 after Theorem 6.8 in [7]). In the general case in [7], Section 4, we described a construction leading to a symmetric subalgebra $\mathfrak{g}_1{}^0 = \mathfrak{k}_1 + \mathfrak{p}_1{}^0$ of $\mathfrak{g}^0$ which is of tube type. In Section 4 of the present paper we shall define certain subalgebras $\mathfrak{g}_\Gamma{}^0$ of $\mathfrak{g}$; the construction leading from $\mathfrak{g}^0$ to $\mathfrak{g}_1{}^0$ can also be performed for $\mathfrak{g}_\Gamma{}^0$, and gives rise to subalgebras $g_{\Gamma,1}{}^0 = k_{\Gamma,1} + p_{\Gamma,1}{}^0$. All these objects will be precisely defined as they occur; here we only wanted to point out the reason for our later notations.

In [7] we determined the Bergman-Šilov boundary S of the bounded symmetric domain D. In the present paper we make a more detailed study of the boundary of D, and will show that it is a union of boundary components, which we describe explicitly. This notion was introduced Pjateckiĭ-Šapiro [9] and is defined as follows: A subset F of the boundary ∂D of D is a *boundary component* if (i) F is locally an analytic set, and (ii) F is minimal with respect to the property that any analytic arc contained in ∂D and having a point in common with F must be entirely contained in F. From our result it follows at once that the Bergman-Šilov boundary of D is exactly the union of all 0-dimensional boundary components.

Index of Notations

Lie algebras and their subsets

$\mathfrak{g}^0$	Lie algebra of G^0, largest connected group of analytic automorphisms of the hermitian symmetric space $M = G^0/K$ of noncompact type
$\mathfrak{k}$	Lie algebra of K
$\mathfrak{p}^0$	(-1)-eigenspace of the symmetry σ on $\mathfrak{g}^0$

$\mathfrak{g}$	$\mathfrak{k}+\mathfrak{p}$, $\mathfrak{p}=i\mathfrak{p}^0$. $\mathfrak{g}$ is the Lie algebra of G where $M^*=G/K$ is the compact dual of M
$\mathfrak{z}$	center of $\mathfrak{k}$
$\mathfrak{h}$	Cartan subalgebra of $\mathfrak{k}$, thus also of $\mathfrak{g}^0$ and $\mathfrak{g}$
$\mathfrak{a}^0$	real span of all $X_\alpha{}^0$ $(\alpha\in\Delta)$, Cartan subalgebra of $(\mathfrak{g}^0,\mathfrak{k})$
$\mathfrak{h}^-;\mathfrak{h}^+$	$[\mathfrak{a}^0, J\mathfrak{a}^0]$; orthogonal complement of $\mathfrak{h}^-$ in $\mathfrak{h}$
$\mathfrak{t}$	$\mathfrak{h}^+ + \mathfrak{a}^0$, maximally split Cartan subalgebra of $\mathfrak{g}^0$
$\mathfrak{g}^C$	complexification of $\mathfrak{g}$, thus also of $\mathfrak{g}^0$

If $\mathfrak{v}$ is a real linear subspace of $\mathfrak{g}^C$:

$\mathfrak{v}^C$	complex span of $\mathfrak{v}$ in $\mathfrak{g}^C$
$\mathfrak{v}^+;\mathfrak{v}^-$	complex span of all positive, or all negative, root vectors in $\mathfrak{v}^C$, except where $\mathfrak{v}=\mathfrak{h},\mathfrak{n}_1{}^\Gamma,\mathfrak{n}_2{}^\Gamma,\mathfrak{n}^\Gamma,\mathfrak{r}_2{}^\Gamma$ or $\mathfrak{r}^\Gamma$.
$\mathfrak{v}^0$	$\mathfrak{v}^C\cap\mathfrak{g}^0$ for $\mathfrak{v}$ such that $\mathfrak{v}=\sigma(\mathfrak{v})$.

If $\mathfrak{v}^0=\sigma(\mathfrak{v}^0)\subset\mathfrak{g}^0$ then $\mathfrak{v}$ denotes $\mathfrak{v}^{0C}\cap\mathfrak{g}$.

If $\alpha\in\Delta$, then $\mathfrak{g}_\alpha$ is the real subalgebra spanned by iH_α, X_α and Y_α.

If Γ is an arbitrary subset of Δ:

$\mathfrak{g}_\Gamma{}^C$	derived algebra of $\mathfrak{h}^C+\sum_{\beta\perp\Delta-\Gamma}E_\beta\cdot C$.
$\mathfrak{k}_\Gamma;\mathfrak{p}_\Gamma;\mathfrak{a}_\Gamma;\mathfrak{h}_\Gamma{}^-$	intersection of $\mathfrak{g}_\Gamma$ with $\mathfrak{k}$, $\mathfrak{p}$, $\mathfrak{a}$, $\mathfrak{h}^-$
$\mathfrak{p}_{\Gamma,1};\mathfrak{k}_{\Gamma,1};\mathfrak{g}_{\Gamma,1}$	$(+1)$-eigenspace of $\tau_\Gamma{}^2$ on $\mathfrak{p}_\Gamma$; $[\mathfrak{p}_{\Gamma,1},\mathfrak{p}_{\Gamma,1}]$; $\mathfrak{k}_{\Gamma,1}+\mathfrak{p}_{\Gamma,1}$
$\mathfrak{l}_{\Gamma,1};\mathfrak{q}_{\Gamma,1};\mathfrak{k}_{\Gamma,1}{}^*$	(±1)-eigenspaces of τ_Γ on $\mathfrak{k}_{\Gamma,1}$; $\mathfrak{l}_{\Gamma,1}+i\mathfrak{q}_{\Gamma,1}$
$\mathfrak{g}^\Gamma;\mathfrak{k}^\Gamma;\mathfrak{p}^\Gamma$	$(+1)$-eigenspace of $\tau_{\Delta-\Gamma}{}^2$ on $\mathfrak{g}$; $\mathfrak{g}^\Gamma\cap\mathfrak{k}$; $\mathfrak{g}^\Gamma\cap\mathfrak{p}$
$\mathfrak{k}_1{}^\Gamma;\mathfrak{p}_1{}^\Gamma;\mathfrak{g}_1{}^\Gamma$	$[\mathfrak{p}^\Gamma,\mathfrak{p}^\Gamma]$; $\mathfrak{p}^\Gamma$; $\mathfrak{k}_1{}^\Gamma+\mathfrak{p}_1{}^\Gamma$
$\mathfrak{l}_1{}^\Gamma;\mathfrak{q}_1{}^\Gamma;\mathfrak{k}_1{}^{\Gamma*}$	(±1)-eigenspaces of $\tau_{\Delta-\Gamma}$ on $\mathfrak{k}_1{}^\Gamma$; $\mathfrak{l}_1{}^\Gamma+i\mathfrak{q}_1{}^\Gamma$
$\mathfrak{l}_2{}^\Gamma$	centralizer of $\mathfrak{g}_1{}^\Gamma$ in $\mathfrak{g}^\Gamma$
$\mathfrak{q}_2{}^\Gamma;\mathfrak{p}_2{}^\Gamma$	(-1)-eigenspace of $\tau_{\Delta-\Gamma}{}^2$ on $\mathfrak{k}$; on $\mathfrak{p}$
$\mathfrak{l}^\Gamma;\mathfrak{q}^\Gamma;\mathfrak{k}^{\Gamma*}$	$\mathfrak{l}_1{}^\Gamma+\mathfrak{l}_2{}^\Gamma$; $\mathfrak{q}_1{}^\Gamma+\mathfrak{q}_2{}^\Gamma$; $\mathfrak{l}^\Gamma+i\mathfrak{q}_1{}^\Gamma$
$\mathfrak{r}_2{}^{\Gamma\pm};\mathfrak{r}^{\Gamma\pm}$	$\mathfrak{q}_2{}^{\Gamma\mp}+\mathfrak{p}_2{}^{\Gamma\pm}$; $\mathfrak{p}_{\Delta-\Gamma,1}{}^\pm+\mathfrak{r}_2{}^{\Gamma\pm}$
$\mathfrak{n}_2{}^{\Gamma\pm};\mathfrak{n}_1{}^{\Gamma\pm};\mathfrak{n}^{\Gamma\pm}$	$\mathfrak{r}_2{}^{\Gamma\pm}\cap\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$; $\mathfrak{p}_{\Delta-\Gamma,1}{}^\pm\cap\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$; $\mathfrak{n}_1{}^{\Gamma\pm}+\mathfrak{n}_2{}^{\Gamma\pm}$
$\mathfrak{b}^\Gamma$	Lie algebra of B^Γ
c_Γ	$K_{\Delta-\Gamma,1}{}^*(-i\mathfrak{o}^\Gamma)$

Mappings

σ	symmetry of $\mathfrak{g}$ or of $\mathfrak{g}^0$
$\nu;\nu^0$	complex conjugation of $\mathfrak{g}^C$ over $\mathfrak{g}$; over $\mathfrak{g}^0$
J	$\mathrm{ad}(Z)$
$\tau;\tau_\Gamma$	$\mathrm{ad}(c)^2$; $\mathrm{ad}(c_\Gamma)^2$
ξ	Harish-Chandra's map $\mathfrak{p}^-\to M^*$ given by $E\to\exp(E)(x)$

Subgroups and submanifolds

G^C	adjoint group of $\mathfrak{g}^C$
Capital roman letter corresponding to a small german letter	corresponding analytic subgroup of G^C, with the following exceptions
$L_1{}^\Gamma; L^\Gamma; E^\Gamma$	isotropy subgroup at x^Γ of $K_1{}^\Gamma$; of K; of G^0
B^Γ	subgroup of G^0 preserving the set $c_{\Delta-\Gamma}M_\Gamma$
$M_\Gamma; M_{\Gamma,1}; M^\Gamma$	submanifolds $G_\Gamma{}^0(x); G_{\Gamma,1}{}^0(x); G^{\Gamma 0}(x)$ of M
$M_\Gamma{}^*; M_{\Gamma,1}{}^*; M^{\Gamma *}$	submanifolds $G_\Gamma(x); G_{\Gamma,1}(x); G^\Gamma(x)$ of M^*
$D; D_\Gamma$	$\xi^{-1}(M); D \cap \mathfrak{p}_\Gamma{}^-$
$S^\Gamma; S_D{}^\Gamma$	set of boundary components of type Γ of M; of D
$U^\Gamma; U_D{}^\Gamma$	union of boundary components of type Γ of M; of D
$\check{S}; \check{S}_D$	Bergman-Šilov boundary S^ϕ of M in M^*; $S_D{}^\phi$ of D in $\mathfrak{p}^-$

Group, algebra and manifold elements

$H_\alpha, E_\alpha, \cdots$	standard basis elements of $\mathfrak{g}^C$
Δ	maximal set of strongly orthogonal noncompact positive roots
$X_\alpha{}^0; X_\alpha$	$E_\alpha + E_{-\alpha} \in \mathfrak{p}^0; iX_\alpha{}^0 \in \mathfrak{p}$
$Y_\alpha{}^0; Y_\alpha$	$-i(E_\alpha - E_{-\alpha}) \in \mathfrak{p}^0; iY_\alpha{}^0 \in \mathfrak{p}$
Z	element of $\mathfrak{z}$ such that $\mathrm{ad}(Z)E = \pm iE$ for $E \in \mathfrak{p}^\pm$
$Z^0; Z'$	$-(i/2)\sum_{\alpha\in\Delta} H_\alpha; Z - Z^0$
$X_\Gamma{}^0; Y_\Gamma{}^0; Z_\Gamma{}^0$	$\sum_{\alpha\in\Gamma} X_\alpha{}^0; \sum_{\alpha\in\Gamma} Y_\alpha{}^0; -i/2 \sum_{\alpha\in\Gamma} H_\alpha$
$X_\Gamma; Y_\Gamma$	$iX_\Gamma{}^0; iY_\Gamma{}^0$
$c_\alpha(\alpha\in\Delta); c; c_\Gamma$	$\exp((\pi/4)X_\alpha) \in G; \prod_{\alpha\in\Delta} c_\alpha; \prod_{\alpha\in\Gamma} c_\alpha$
x	identity coset in $M^* = G^C/K^C P^+$
x^Γ	$c_{\Delta-\Gamma}(x)$
$\mathfrak{o}; \mathfrak{o}^\Gamma$	zero elements $\xi^{-1}(x)$ of $\mathfrak{p}^-$; $\xi^{-1}(x^\Gamma)$

3. A theorem on real parabolic groups. We will classify a certain family of real parabolic subgroups of the Lie groups which are the connected groups of analytic automorphisms of the bounded symmetric domains. In Corollary 6.9 it will be seen that those parabolic subgroups are just the stability groups of the various boundary components. We will also need the notion of parabolic group in our proof of Theorem 6.8.

The goal of this section is Theorem 3.4, which resulted from a con-

versation between J. Tits and one of the authors. All the other results of this section are special cases of theorems of A. Borel and J. Tits [2] on linear algebraic groups.

3.1. *Parabolic subgroups of complex Lie groups.* Let E be a complex connected Lie group. Then the maximal solvable subgroups of E are all closed, complex, connected and conjugate; they are called the *Borel subgroups* of E and their Lie algebras are the *Borel subalgebras* of $\mathfrak{e}$. If a complex Lie subgroup of E contains a Borel subgroup, then it and its Lie algebra are called *parabolic.* Every parabolic subgroup $F \subset E$ is connected, for every component of F contains an element which normalizes a Borel subgroup B of F_0 (and thus of E), and it follows that this component must be F_0 because it contains an element which centralizes a Cartan subgroup of E which lies in B. Similarly every parabolic subgroup $F \subset E$ is its own normalizer. As every Borel subgroup of E contains the radical of E, we may pass to a quotient and restrict our study to the case where E is semisimple.

Let E be a connected complex semisimple Lie group. Choose a Cartan subalgebra $\mathfrak{c}$ of the Lie algebra $\mathfrak{e}$, let Δ denote the root system of $\mathfrak{e}$ relative to $\mathfrak{c}$, and choose a simple system Ψ of roots. If $\mathfrak{e}_\lambda$ denotes the root space for $\lambda \in \Delta$, and if Δ^+ denotes the set of positive roots, then our choices amount to the choice of the Borel subalgebra

$$\mathfrak{b} = \mathfrak{c} + \sum_{\lambda \in \Delta^+} \mathfrak{e}_\lambda$$

of $\mathfrak{e}$. Now let $\Phi \subset \Psi$, and define

$$\Phi^+ = \{\lambda \in \Delta : \lambda = \sum_{\alpha \in \Psi} a_\alpha \alpha \text{ with } a_\alpha > 0 \text{ for some } \alpha \in \Phi\},$$

$$\Phi^0 = \{\lambda \in \Delta : \lambda = \sum_{\alpha \in \Psi} a_\alpha \alpha \text{ with } a_\alpha = 0 \text{ for every } \alpha \in \Phi\}, \text{ and}$$

$$\Phi^* = \Phi^0 \cup \Phi^+ = \{\lambda \in \Delta : \lambda = \sum_{\alpha \in \Psi} a_\alpha \alpha \text{ with } a_\alpha \geqq 0 \text{ for every } \alpha \in \Phi\}.$$

Then

$$\mathfrak{f}_\Phi = \mathfrak{c} + \sum_{\lambda \in \Phi^*} \mathfrak{e}_\lambda$$

is a parabolic subalgebra of $\mathfrak{e}$ which contains $\mathfrak{b}$. $\mathfrak{f}_\Psi = \mathfrak{b}$, $\mathfrak{f}_\phi = \mathfrak{e}$, and $\mathfrak{f}_\Sigma \subset \mathfrak{f}_\Gamma$ for $\Gamma \subset \Sigma \subset \Psi$. Conversely, let $\mathfrak{f}$ be a parabolic subalgebra of $\mathfrak{e}$ which contains $\mathfrak{b}$, and define

$$\Phi = \{\alpha \in \Psi : a_\alpha \geqq 0 \text{ whenever } \mathfrak{e}_\lambda \subset \mathfrak{f} \text{ with } \lambda = \sum_{\beta \in \Psi} a_\beta \beta\}.$$

Then it is routine to check that $\mathfrak{f} = \mathfrak{f}_\Phi$. Now we can specify a conjugacy

class of parabolic subgroups of E by marking the elements of $\Psi - \Phi$ on the Dynkin diagram of $\mathfrak{e}$, where $\mathfrak{f}_\Phi$ is the Lie algebra of an element of this class.

Retain the notation just above, and define (for every subset $\Phi \subset \Psi$)

$$\mathfrak{c}_\Phi = \bigcap_{\alpha \in \Psi - \Phi} (\text{kernel of } \alpha),$$

$$\mathfrak{r}_\Phi = \mathfrak{c} + \sum_{\lambda \in \Phi^0} \mathfrak{e}_\lambda, \text{ and}$$

$$\mathfrak{u}_\Phi = \sum_{\lambda \in \Phi^+} \mathfrak{e}_\lambda.$$

Then $\mathfrak{f}_\Phi = \mathfrak{r}_\Phi + \mathfrak{u}_\Phi$ (semidirect sum) and is the normalizer of $\mathfrak{u}_\Phi$ in $\mathfrak{e}$. $\mathfrak{u}_\Phi$ is nilpotent, $\mathfrak{r}_\Phi$ is reductive in $\mathfrak{e}$ because it is the centralizer of $\mathfrak{c}_\Phi$ in $\mathfrak{e}$, and $\mathfrak{c}_\Phi$ is the center of $\mathfrak{r}_\Phi$. Let R_Φ, U_Φ and F_Φ denote the analytic subgroups of E for the subalgebras $\mathfrak{r}_\Phi$, $\mathfrak{u}_\Phi$ and $\mathfrak{f}_\Phi$ of $\mathfrak{e}$. Then $F_\Phi = R_\Phi \cdot U_\Phi$ semidirect product. We may view E as a linear algebraic group because it is complex, connected and semisimple, and then this semidirect product decomposition of P_Φ is the Chevalley decomposition into reductive and unipotent parts.

3.2. *Parabolic subgroups of real Lie groups.* Let E' be a connected semisimple real Lie group embedded in its complexification. In other words there is a complex connected semisimple Lie group E and a real form $\mathfrak{e}'$ of the complex Lie algebra $\mathfrak{e}$ such that E' is the real analytic subgroup of E with Lie algebra $\mathfrak{e}'$. We will say that a subgroup $F' \subset E'$ is a *parabolic subgroup* of E' if there exists a parabolic subgroup $F \subset E$ such that (i) $F' = E' \cap F$ and (ii) $\mathfrak{f}$ is the complexification of $\mathfrak{f}'$. If F' is a parabolic subgroup of E' and F_0' denotes the identity component, then any element $f \in F'$ can be altered by an element of F_0' to centralize a Cartan subalgebra $\mathfrak{c}'$ of $\mathfrak{e}'$ contained in $\mathfrak{f}'$, and it follows that $F' = (C \cap E') \cdot F_0'$ where $C = \exp(\mathfrak{c})$ and $\mathfrak{c}$ is the Cartan subalgebra of $\mathfrak{e}$ which is the complexification of $\mathfrak{c}'$.

Let $\mathfrak{c}'$ be a Cartan subalgebra of $\mathfrak{e}'$. Then there is a canonical decomposition $\mathfrak{c}' = \mathfrak{c}_t + \mathfrak{c}_v$ where the roots are real valued on $\mathfrak{c}_v$ and take pure imaginary values on $\mathfrak{c}_t$. To obtain this decomposition, consider the Cartan subgroup $C' = \exp(\mathfrak{c}') \subset E'$. C' has a unique maximal compact subgroup C_t, and $\mathfrak{c}_t$ is the corresponding subalgebra; it is clear that the roots take pure imaginary values on $\mathfrak{c}_t$. Let $\mathfrak{c}_v$ be the orthogonal complement of $\mathfrak{c}_t$ in $\mathfrak{c}'$ under the Killing form. Then $C_v = \exp(\mathfrak{c}_v)$ is a vector subgroup of C'. If $\mathfrak{c}_v$ and and C_v were not diagonable in $\mathrm{ad}(\mathfrak{e}')$ on $\mathfrak{e}'$, C_t would not be maximal. Thus the roots are real valued on $\mathfrak{c}_v$. C_t (resp. $\mathfrak{c}_t$) and C_v (resp. $\mathfrak{c}_v$) are the *totally non-split* and the *split* parts of C' (resp. $\mathfrak{c}'$). C' and $\mathfrak{c}'$ are *maximally*

split if $\dim . \mathfrak{c}_v$ is maximal among the dimensions of the split parts of the Cartan subalgebra of $\mathfrak{e}'$.

LEMMA. *The parabolic subgroups of E' are just the subgroups $F' = F \cap E'$ for which there exist* (a) *a maximally split Cartan subalgebra $\mathfrak{c}'$ of $\mathfrak{e}'$,* (b) *a system Ψ of simple roots of $\mathfrak{e}$ for $\mathfrak{c} = \mathfrak{c}'^C$ and* (c) *a subset $\Phi \subset \Psi$, such that* (i) *$\mathfrak{c}'_\Phi$ is a real form of $\mathfrak{c}_\Phi$,* (ii) *$\mathfrak{c}'_\Phi$ and its split part $\mathfrak{c}'_\Phi \cap \mathfrak{c}_v$ have the same centralizer in $\mathfrak{e}$,* (iii) *$\mathfrak{f}_\Phi$ is the sum of the non negative weight spaces of* $\mathrm{ad}(\mathfrak{c}'_\Phi \cap \mathfrak{c}_v)$ *on $\mathfrak{e}$, and* (iv) *$F = F_\Phi$.*

Proof. Let ν be conjugation of $\mathfrak{e}$ over $\mathfrak{e}'$. Given $\mathfrak{c}'$, Ψ and Φ satisfying (i)-(iv), we recall that the roots of $\mathfrak{e}$ are real-valued on $\mathfrak{c}_v$. Thus every weight space of $\mathrm{ad}(\mathfrak{c}'_\Phi \cap \mathfrak{c}_v)$ on $\mathfrak{e}$ is stable under ν. Now (iii) says that $\mathfrak{f}_\Phi = (\mathfrak{f}_\Phi \cap \mathfrak{e}')^C$, so (iv) tells us that $F' = F \cap E'$ is parabolic in E'.

Let F' be a parabolic subgroup of E'. Then $F' = F \cap E'$ for some parabolic subgroup F of E, and $\nu(\mathfrak{f}) = \mathfrak{f}$. ν preserves the maximal nilpotent normal subalgebra $\mathfrak{u}$ of $\mathfrak{f}$ and we choose a ν-invariant reductive complement $\mathfrak{r}$. If $\mathfrak{c}_*$ denotes the center of $\mathfrak{r}$, then $\nu(\mathfrak{c}_*) = \mathfrak{c}_*$ so $\mathfrak{c}'_* = \mathfrak{c}_* \cap \mathfrak{e}'$ is a real form of $\mathfrak{c}_*$. There is a lexicographic ordering on the dual space of the real form $\mathfrak{c}_{*v} + i\mathfrak{c}_{*t}$ of $\mathfrak{c}_*$, such that $\mathfrak{u}$ is the sum of the positive weight spaces. We extend $\mathfrak{c}_*$ to a Cartan subalgebra $\mathfrak{c}$ of $\mathfrak{e}$ for which $\mathfrak{c}' = \mathfrak{c} \cap \mathfrak{e}'$ is a real form, in such a manner as to maximize the dimension of the split part of $\mathfrak{c}'$; we then extend the ordering of weights on $\mathfrak{c}_*$ to an ordering of the $\mathfrak{c}$-roots of $\mathfrak{e}$. Let Ψ be the corresponding system of simple roots. Now $F = F_\Phi$ for some subset $\Phi \subset \Psi$.

We must check that the $\mathfrak{c}'$, Ψ and Φ just constructed satisfy the conditions (i), (ii) and (iii) and that $\mathfrak{c}'$ is maximally split. Condition (i) is immediate because $\mathfrak{c}_* = \mathfrak{c}_\Phi$ from the construction of complex parabolic algebras. The split part $\mathfrak{c}_{*v} = \mathfrak{c}'_\Phi \cap \mathfrak{c}_v$ of $\mathfrak{c}'_\Phi$ is nonzero because the sum of the elements of Φ induces a positive linear functional on it. Now $\mathfrak{e} \supsetneqq \mathfrak{r}_* \supset \mathfrak{r}$ where $\mathfrak{r}_*$ is the centralizer of $\mathfrak{c}_{*v}$ and $\mathfrak{r}$, the reductive part of $\mathfrak{f}$, is the centralizer of $\mathfrak{c}_\Phi$. If $\mathfrak{r}_* \neq \mathfrak{r}$ then $\mathfrak{u} \cap \mathfrak{r}_* \neq 0$, so $\mathfrak{u} \cap \mathfrak{r}_*$ is a nontrivial sum of root spaces. The roots which enter into this sum belong to Φ^+, because $\mathfrak{u} \cap \mathfrak{r}_* \subset \mathfrak{u}$, so their negatives do not appear. Thus the roots which enter the sum must vanish on $\mathfrak{c}_{*t}$, and thus on $\mathfrak{c}'_* = \mathfrak{c}_{*t} + \mathfrak{c}_{*v}$. That is impossible. Now $\mathfrak{r}_* = \mathfrak{r}$ and (ii) is proved. (iii) follows by our ordering of roots so that $\mathfrak{c}_{*v}$ precedes $i\mathfrak{c}_{*t}$. From (ii) we also see that $\mathfrak{r}'$ contains a maximally split Cartan subalgebra of $\mathfrak{e}'$, so the maximality condition in our choice of $\mathfrak{c}$ implies that $\mathfrak{c}'$ is maximally split in $\mathfrak{e}'$. *Q.E.D.*

Lemma 3.2 shows that real parabolic groups are determined by the split parts of the centers of their reductive parts. Let F' be any parabolic subgroup of E'. In the notation of Lemma 3.2, let B' be the group $B \cap E'$ where $\mathfrak{b}$ is the sum of the non negative weight spaces of $\mathrm{ad}(\mathfrak{c}_v)$ on $\mathfrak{e}$. Then B' is a parabolic subgroup of E', $B' \subset F'$, and every parabolic subgroup of E' contains a conjugate of B'. B' is a *minimal parabolic subgroup* of E'. There is an Iwasawa decomposition $E' = K \cdot A \cdot N$ such that $B' = L \cdot A \cdot N$ where $L \subset K$ is the centralizer of A. Furthermore there is a subset $\Sigma \subset \Psi$ defined by $B = F_\Sigma$, such that $\Phi \subset \Sigma$ whenever $F_\Phi \cap E'$ is parabolic in E'.

3.3. *The action of the Galois group.* The Galois group of C over R acts on $\mathfrak{e}$ as the conjugation ν of $\mathfrak{e}$ over $\mathfrak{e}'$. This is a real automorphism of $\mathfrak{e}$ and induces a real automorphism of E. The fixed point set of ν on E is E'. Let F be a parabolic subgroup of E. Now $F \cap E'$ is parabolic in E' if and only if $\nu(F) = F$, and this is equivalent to $\nu(\mathfrak{f}) = \mathfrak{f}$.

Let $\mathfrak{c}$ be a Cartan subalgebra of $\mathfrak{e}$ which is the complexification of a maximality split Cartan subalgebra $\mathfrak{c}'$ of $\mathfrak{e}'$, and let Ψ be a system of simple roots. Then the Galois group $\{1, \nu\}$ acts on Ψ as follows. The subsets of Ψ are in one-one correspondence with the conjugacy classes of parabolic subgroups of E, a subset Φ corresponding to the class of F_Φ. Given Φ, $\nu(F_\Phi)$ is conjugate to some F_Σ and we define $\Sigma = \Phi^\nu$. This transformation on the subsets of Ψ is induced by its restriction to the one-point subsets, so ν acts on Ψ. If $E' \cap F_\Phi$ is parabolic in E', then $\nu(F_\Phi) = F_\Phi$, and so $\Phi^\nu = \Phi$. The converse is:

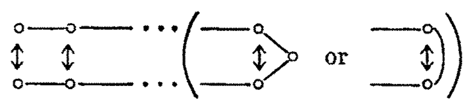
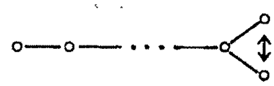
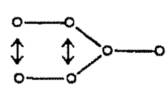
LEMMA. *Let Ψ be a system of simple roots of $\mathfrak{e}$ for the complexification of a maximally split Cartan subalgebra $\mathfrak{c}'$ of $\mathfrak{e}'$, and let Σ be the subset of Ψ such that $E' \cap F_\Sigma$ is a minimal parabolic subgroup of E'. Then the parabolic subgroups of E' are just the conjugates of the groups $E' \cap F_\Phi$ for which $\Phi \subset \Sigma$ and $\Phi^\nu = \Phi$.*

Proof. The remark above and the results of § 3.2 show that $\Phi \subset \Sigma$ and $\Phi^\nu = \Phi$ in case $E' \cap F_\Phi$ is parabolic in E'.

Let $\Phi \subset \Sigma$ and $\Phi^\nu = \Phi$; we will check that $E' \cap F_\Phi$ is parabolic in E'. If two parabolic subgroups of E are conjugate and contain the same Borel subgroup, then they must be the same. Now if two parabolic subgroups of E are conjugate and contain F_Σ they must be the same, for F_Σ contains a Borel subgroup. $\nu(F_\Sigma) = F_\Sigma$, and $F_\Sigma \subset F_\Phi$ because $\Phi \subset \Sigma$; thus F_Φ and $\nu(F_\Phi)$ contain F_Σ. $\Phi = \Phi^\nu$ says that F_Φ is conjugate to $\nu(F_\Phi)$. Thus $F_\Phi = \nu(F_\Phi)$. This proves that $E' \cap F_\Phi$ is parabolic in E'. *Q.E.D.*

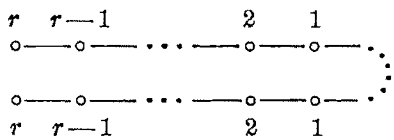
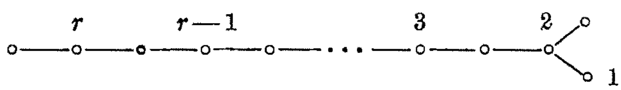
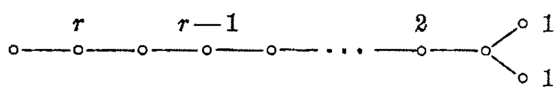
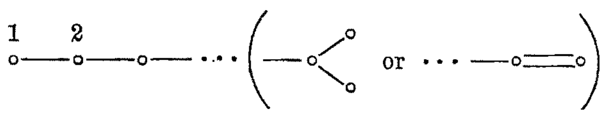
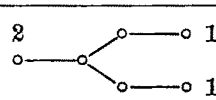
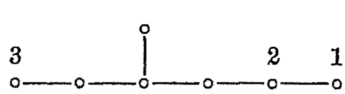
In order to apply the Lemma, one must know the action of the Galois group on Ψ.

Complement to Lemma. *Let $\mathfrak{g}'$ be a real simple Lie algebra with simple complexification $\mathfrak{g}$, let ν be the nontrivial element of the Galois group of C over R, and let Ψ be a system of simple roots of $\mathfrak{g}$ for the complexification of a maximally split Cartan subalgebra of $\mathfrak{g}'$. Then the action of ν on Ψ is trivial except in the following cases.*

$\mathfrak{g}'$	action of ν on Dynkin diagram
$\mathfrak{su}^k(n)$	
$\mathfrak{so}^*(4n+2)$ *or* $\mathfrak{so}^{2k}(4n+2)$ *or* $\mathfrak{so}^{2k+1}(4n)$	
$\mathfrak{e}_{6(-14)}$ *or* $\mathfrak{e}_{6(+2)}$	

Proof. $\mathfrak{g}'$ has a maximal compactly embedded subalgebra $\mathfrak{k}$ such that, if σ denotes the symmetry of the symmetric pair $(\mathfrak{g}', \mathfrak{k})$ and $\mathfrak{g}' = \mathfrak{k} + \mathfrak{p}$ is the Cartan decomposition, $\mathfrak{k} \cap \mathfrak{c}' = \mathfrak{c}_t$ and $\mathfrak{p} \cap \mathfrak{c}' = \mathfrak{c}_v$. The root vectors are in $i\mathfrak{c}_t + \mathfrak{c}_v$, so ν is -1 on their $\mathfrak{c}_t$ projections and is $+1$ on their $\mathfrak{c}_v$ projections. σ is complex linear on $\mathfrak{g}$, $+1$ on $\mathfrak{k}$ and -1 on $\mathfrak{p}$; thus $\sigma\nu$ sends each root vector to its negative. Thus (i) if σ is an inner automorphism, then ν is trivial on Ψ precisely in case $-I$ is in the Weyl group of $\mathfrak{g}$, and (ii) if σ is an outer automorphism, then ν is trivial on Ψ precisely in case $-I$ is not in the Weyl group. As $-I$ is in the Weyl group in all cases except $\mathfrak{g} = A_n$ $(u > 1)$, D_{2n+1} $(n > 1)$, or E_6 [12, Theorem 4.1], the result follows from the classification of the real simple Lie algebras. *Q.E.D.*

3.4. THEOREM. *Let M be an irreducible hermitian symmetric space of noncompact type; let G^0 be the connected group of analytic automorphisms of M, embedded in its complexification G^C; let $\mathfrak{u}$ be a maximally split Cartan subalgebra of $\mathfrak{g}^0$ which is preserved by the symmetry at a point $x \in M$, and let Ψ be a system of simple roots of $\mathfrak{g}^C$ for $\mathfrak{u}^C$. If r is the rank of M, then there is a unique sequence $\{\Phi_0, \Phi_1, \cdots, \Phi_r\}$ of subsets of Ψ such that* (i) *$F_{\Phi_i}{}^0 = G^0 \cap F_{\Phi_i}$ is a parabolic subgroup of G^0,* (ii) *the reductive part of $F_{\Phi_i}{}^0$ has a simple normal subgroup $G_i{}^0$ such that $G_i{}^0(x)$ is a hermitian symmetric subspace of rank i in M, and* (iii) *the $G_i{}^0$ can be chosen so that $G_0{}^0 \subset G_1{}^0 \subset \cdots \subset G_r{}^0$. Furthermore, $G_r{}^0 = F_{\Phi_r}{}^0 = G^0$ and Φ_i consists of the elements of Ψ numbered $\{i+1, \cdots, r-1, r\}$ in the chart below.*

$\mathfrak{g}^0$	*Dynkin diagram*
$\mathfrak{su}^r(n)$ $(2r \leqq n)$	
$\mathfrak{so}^*(4r)$	
$\mathfrak{so}^*(4r+2)$	
$\mathfrak{so}^2(n+2)$ $(r=2)$	
$\mathfrak{sp}(r, R)$	
$\mathfrak{e}_{6(-14)}$ $(r=2)$	
$\mathfrak{e}_{7(-25)}$ $(r=3)$	

Proof. To satisfy (ii) and (iii) we must have $\Phi_r \subsetneqq \Phi_{r-1} \subsetneqq \cdots \subsetneqq \Phi_0$. Let $\Sigma \subset \Psi$ so that $G^0 \cap F_\Sigma$ is a minimal parabolic subgroup of G^0. Then $\Phi_0 \subset \Sigma$ by (i). The Galois group has exactly r orbits on Σ, and each Φ_i is a union of orbits by (iii). It follows that we can number the orbits as $\Sigma_1, \cdots, \Sigma_r$ so that $\Phi_i = \Sigma_{i+1} \cup \cdots \cup \Sigma_r$; $\Phi_r \neq \phi$ and $\Phi_0 = \Sigma$. Now a case by case check, using the fact that $\mathfrak{g}_i{}^0$ must be one of the algebras listed for the $\mathfrak{g}^0$ with $i = r$, and using the fact that the Dynkin diagram of $\mathfrak{g}_i{}^0$ must be a connected component of the complement of $\Sigma_{i+1} \cup \cdots \cup \Sigma_r$ in the diagram of $\mathfrak{h}^0$, shows that (ii) and (iii) imply that Σ_i must consist of the points numbered i in the Dynkin diagram, and that then (i), (ii) and (iii) are satisfied. The Theorem follows. *Q.E.D.*

4. The boundary components of a bounded symmetric domain. Let D be a bounded symmetric domain embedded in $\mathfrak{p}^-$ as described in § 2. Retain the notation of § 2, and let Γ be an arbitrary subset of the maximal set Δ of strongly orthogonal noncompact roots. We will see that every $\Gamma \subsetneqq \Delta$ corresponds to a certain boundary component of D, the empty set ϕ corresponding to a point of the Bergman-Šilov boundary, and Δ corresponding to D. It turns out that two subsets of Δ give analytically equivalent boundary components if and only if they contain the same number of roots from each irreducible factor of M.

4. 1. We have $c_\alpha = \exp(\pi/4) X_\alpha \in G$ for every $\alpha \in \Delta$, and the Cayley transform on M was defined to be $c = \prod_{\alpha \in \Delta} c_\alpha$. We now define *partial Cayley transforms* by

$$c_\Gamma = \prod_{\alpha \in \Gamma} c_\alpha, \ \Gamma \subset \Delta,$$

so $c_\Delta = c$ and $c_\phi = 1$. Similarly we define

$$X_\Gamma{}^0 = \sum_{\alpha \in \Gamma} X_\alpha \text{ and } X_\Gamma = i X_\Gamma{}^0,$$

$$Y_\Gamma{}^0 = \sum_{\alpha \in \Gamma} Y_\alpha{}^0 \text{ and } Y_\Gamma = i Y_\Gamma{}^0, \text{ and}$$

$$Z_\Gamma{}^0 = -\tfrac{1}{2} \sum_{\alpha \in \Gamma} H_\alpha \text{ and } \mathfrak{h}_\Gamma{}^- = \sum_{\alpha \in \Gamma} i H_\alpha \cdot R.$$

By definition of H_β the centralizer of $\mathfrak{h}_{\Delta-\Gamma}{}^-$ in $\mathfrak{g}^C$ is

$$\mathfrak{h}^C + \sum_{\beta \perp \Delta-\Gamma} E_\beta \cdot C. \tag{4.1.1}$$

The centralizer of $\sum_{\alpha \in \Delta-\Gamma} \mathfrak{g}_\alpha$ in $\mathfrak{g}^C$ is

$$\mathfrak{h}^{+C} + \mathfrak{h}_\Gamma{}^{-C} + \sum_{\beta \perp \Delta-\Gamma} E_\beta \cdot C. \tag{4.1.2}$$

For this, it suffices to show that $\beta \perp \Delta - \Gamma$ implies $[E_\beta, E_{\pm\alpha}] = 0$ whenever $\alpha \in \Delta - \Gamma$. Here we may assume $\beta > 0$; then the assertion is trivial for β noncompact and known [4, Lemma 13] for β compact.

The algebras (4.1.1) and (4.1.2) are reductive and have the same derived algebra. We denote this derived algebra by $\mathfrak{g}_\Gamma{}^C$, and $\mathfrak{g}_\Gamma = \mathfrak{g} \cap \mathfrak{g}_\Gamma{}^C$ and $\mathfrak{g}_\Gamma{}^0 = \mathfrak{g}_\Gamma{}^0 \cap \mathfrak{g}_\Gamma{}^C$ are real forms of $\mathfrak{g}_\Gamma{}^C$; they are semisimple. We have $\mathfrak{g}_\Gamma = \mathfrak{k}_\Gamma + \mathfrak{p}_\Gamma$ and $\mathfrak{g}_\Gamma{}^0 = \mathfrak{k}_\Gamma + \mathfrak{p}_\Gamma{}^0$ where $\mathfrak{k}_\Gamma = \mathfrak{g}_\Gamma \cap \mathfrak{k} = \mathfrak{g}_\Gamma{}^0 \cap \mathfrak{k}$, $\mathfrak{p}_\Gamma = \mathfrak{g}_\Gamma \cap \mathfrak{p}$ $\mathfrak{p}_\Gamma{}^0 = \mathfrak{g}_\Gamma{}^0 \cap \mathfrak{p}^0$, and $\mathfrak{p}_\Gamma{}^0 = i\mathfrak{p}_\Gamma$. $\mathfrak{p}_\Gamma{}^{\pm}$ denotes $\mathfrak{p}_\Gamma{}^C \cap \mathfrak{p}^{\pm}$. G_Γ and $G_\Gamma{}^0$ denote the respective analytic subgroups of G^C with Lie algebras $\mathfrak{g}_\Gamma$ and $\mathfrak{g}_\Gamma{}^0$, and K_Γ denotes their common intersection with K.

Further, we define $M_\Gamma = G_\Gamma{}^0(x)$, $M_\Gamma{}^* = G_\Gamma(x)$, and $x^\Gamma = c_{\Delta-\Gamma}(x)$. In the special case where Γ is empty, M_Γ and $M_\Gamma{}^*$ are just $\{x\}$ and x^Γ is the point $c(x)$ on the Bergman-Šilov boundary. More generally, we will eventually see that $c_{\Delta-\Gamma}(M_\Gamma) = G_\Gamma{}^0(x^\Gamma)$ and is a typical boundary component of M in M^*.

Finally define $D_\Gamma = D \cap \mathfrak{p}_\Gamma{}^-$ and $\mathfrak{o}^\Gamma = \xi^{-1}(x^\Gamma)$; $\mathfrak{o}$ will denote the origin, $\mathfrak{o} = \mathfrak{o}^\Delta = 0$, of $\mathfrak{p}^-$.

4.2. Lemma. *M_Γ is a complex totally geodesic submanifold of M, thus being a sub hermitian symmetric space of M; the same is true for $M_\Gamma{}^*$ in M^*, and $M_\Gamma \subset M_\Gamma{}^*$ is the Borel embedding. Γ is a maximal set of strongly orthogonal noncompact roots of $\mathfrak{g}_\Gamma{}^0$, $\{X_\alpha{}^0\}_{\alpha\in\Gamma}$ spans a Cartan subalgebra $\mathfrak{a}_\Gamma{}^0 = \mathfrak{a}^0 \cap \mathfrak{g}_\Gamma{}^0$ of $(\mathfrak{g}_\Gamma{}^0, \mathfrak{k}_\Gamma)$, and c_Γ is the Cayley transform of M_Γ. Let $\xi^{-1}: M \to \mathfrak{p}^-$ be the Harish-Chandra embedding as a bounded domain $D = \xi^{-1}(M)$; then $D_\Gamma = \xi^{-1}(M_\Gamma)$ and $\xi^{-1}: M_\Gamma \to \mathfrak{p}_\Gamma{}^-$ is the Harish-Chandra embedding. $c_{\Delta-\Gamma}(M_\Gamma) \subset \xi(\mathfrak{p}^-)$, $\xi^{-1}c_{\Delta-\Gamma}\xi$ acts on D_Γ by $E \to E + i \sum_{\beta\perp\Delta-\Gamma} E_{-\alpha}$.*

Proof. The algebras 4.1.1 and 4.1.2 are preserved by $\mathrm{ad}(\mathfrak{h})$, and thus by $\mathrm{ad}(Z)$, so $\mathrm{ad}(Z)$ preserves $\mathfrak{g}_\Gamma{}^C$; as $\mathrm{ad}(Z)$ preserves $\mathfrak{g}$ and $\mathfrak{g}^0$, it must preserve $\mathfrak{g}_\Gamma$ and $\mathfrak{g}_\Gamma{}^0$; now $\mathrm{ad}(\exp(tZ))$ preserves G_Γ and $G_\Gamma{}^0$, so $M_\Gamma{}^* \subset M^*$ and $M_\Gamma \subset M$ are sub hermitian symmetric spaces. $M_\Gamma \subset M_\Gamma{}^*$ is the Borel embedding by construction.

The definitions of $\mathfrak{g}_\Gamma{}^0$ and $\mathfrak{g}_{\Delta-\Gamma}{}^0$ give us

$$\sum_{\alpha\in\Gamma} X_\alpha{}^0 \cdot R \subset \mathfrak{a}_\Gamma{}^0, \quad \Big(\sum_{\alpha\in\Gamma} X_\alpha{}^0 \cdot R\Big) \cap \mathfrak{a}_{\Delta-\Gamma}{}^0 = 0, \text{ and } \sum_{\alpha\in\Delta-\Gamma} X_\alpha{}^0 \cdot R \subset \mathfrak{a}_{\Delta-\Gamma}{}^0;$$

linear independence of the $X_\alpha{}^0$ now shows that $\mathfrak{a}_\Gamma{}^0$ has $\{X_\alpha{}^0\}_{\alpha\in\Gamma}$ for a basis. If $\mathfrak{a}_\Gamma{}^0$ is not a Cartan subalgebra of $(\mathfrak{g}_\Gamma{}^0, \mathfrak{k}_\Gamma)$, then it is properly contained in one, say in $\mathfrak{e}$. $[\mathfrak{e}, \mathfrak{a}_{\Delta-\Gamma}{}^0] = 0$ by definition of $\mathfrak{g}_\Gamma{}^0$, so we have

$$\mathfrak{a}^0 = \mathfrak{a}_\Gamma{}^0 + \mathfrak{a}_{\Delta-\Gamma}{}^0 \subsetneq \mathfrak{e} + \mathfrak{a}_{\Delta-\Gamma}{}^0 \subset \mathfrak{f}$$

where $\mathfrak{f}$ is a Cartan subalgebra of $\mathfrak{g}_\Gamma{}^0$. Then $\dim.\mathfrak{f} = \dim.\mathfrak{a}^0$, which is a contradiction. Now $\mathfrak{a}_\Gamma{}^0$ is a Cartan subalgebra of $(\mathfrak{g}_\Gamma{}^0, \mathfrak{k}_\Gamma)$, and the assertions on Γ and c_Γ follow. It also follows that the Harish-Chandra embedding of M as $D \subset \mathfrak{p}^-$ induces that of M_Γ as $D_\Gamma \subset \mathfrak{p}_\Gamma{}^-$.

As $c_{\Delta-\Gamma}$ commutes with every element of $G_\Gamma{}^0$, the proof of the last statement reduces to proving that $\xi^{-1}c_{\Delta-\Gamma}\xi \colon 0 \to \sum_{\alpha\in\Delta-\Gamma} iE_{-\alpha}$, i.e., that $\xi(i\sum_{\alpha\in\Delta-\Gamma} E_{-\alpha}) = x^\Gamma$; this is a calculation contained in the proof of [7, Lemma 4.2]. *Q.E.D.*

The following is the first step toward relating the M_Γ to the boundary components of D.

4.3. LEMMA. *∂D is the union of all sets of the form*

4.3.1. $$\mathrm{ad}(k)[\xi^{-1}c_{\Delta-\Sigma}M_\Sigma],\ k \in K,\ \Sigma \subsetneqq \Delta.$$

and every boundary component of D is a union of sets of that form.

Proof. $A^0 = \exp(\mathfrak{a}^0)$ consists of transvections of M, so 1 is the only element of $\xi^{-1}A^0\xi$ with a fixed point on D. From the action of the latter on $\mathfrak{a}^-$ [7, Lemma 3.5] it follows that $\partial D \cap \mathfrak{a}^-$ consists of all $\sum_{\alpha\in\Delta} b_\alpha E_{-\alpha}$ with $-1 \leqq b_\alpha \leqq 1$ where at least one $|b_\alpha| = 1$. In particular, every $\sum_{\alpha\in\Sigma} E_{-\alpha} \in \partial D$; as $\xi^{-1}\exp(tZ)\xi$ acts on D and ∂D by unimodular complex scalars, we have $0^\Sigma = i\sum_{\alpha\in\Sigma} E_{-\alpha} \in \partial D$. Applying $G_\Sigma{}^0$, $\xi^{-1}c_{\Delta-\Sigma}M_\Sigma = (\xi^{-1}G_\Sigma{}^0\xi)(0^\Sigma) \subset \partial D$. Thus ∂D contains every set of the form (4.3.1).

We wish to show that ∂D is the union of the sets (4.3.1). As $\partial D = \mathrm{ad}(K)[\partial D \cap i\mathfrak{a}^-]$, it suffices to show that every point of $\partial D \cap i\mathfrak{a}^-$ lies in a set of that form. Every such element has expression $E' = i\sum_{\alpha\in\Delta-\Sigma} \pm E_{-\alpha} + i\sum_{\alpha\in\Sigma} b_\alpha E_{-\alpha}$ where $-1 < b_\alpha < 1$ and $\Sigma \subsetneqq \Delta$; applying an element of $\mathrm{ad}(K \cap \exp \sum_{\alpha\in\Delta-\Sigma} \mathfrak{g}_\alpha)$ we bring it to $E = i\sum_{\alpha\in\Delta-\Sigma} E_{-\alpha} + i\sum_{\alpha\in\Sigma} b_\alpha E_{-\alpha}$, which is in $\xi^{-1}c_{\Delta-\Sigma}M_\Sigma$ by Lemma 4.2. This completes the proof that ∂D is the union of all the sets (4.3.1).

As M_Σ is a hermitian symmetric space of noncompact type by Lemma 4.2, any two of its points can be joined by an analytic arc. It follows that any two points of $\mathrm{ad}(k)[\xi^{-1}c_{\Delta-\Sigma}M_\Sigma]$ can be joined by an analytic arc in ∂D. This completes the proof. *Q.E.D.*

4.4. LEMMA.[3] *The restriction of* $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$ *to* $\mathfrak{p}^C$ *has only the eigen-*

[3] This lemma and a sharpened form (Lemma 6.3) will be used repeatedly. The apparently elaborate notation will be re-introduced and motivated in § 5.4.

values 0, $\pm i$ *and* $\pm i/2$; *the respective eigenspaces are* $\mathfrak{p}_\Gamma{}^C$, *the centralizer* $\mathfrak{p}_{\Delta-\Gamma,1}{}^C$ *of* $c_{\Delta-\Gamma}{}^4$ *in* $\mathfrak{p}_{\Delta-\Gamma}{}^C$, *and the* (-1)*-eigenspace of* $\mathrm{ad}(c_{\Delta-\Gamma})^4$ *in* $\mathfrak{p}^C$.

Proof. Let $\mathfrak{e}$ and $\mathfrak{f}$ be the $+1$ and -1 eigenspaces of $\mathrm{ad}(c_{\Delta-\Gamma})^4$ on $\mathfrak{p}^C$. As $c_{\Delta-\Gamma}$ is a transvection of order 4 or 8 in M^*, so $(c_{\Delta-\Gamma}{}^4)^2 = 1$ and $\sigma(c_{\Delta-\Gamma}{}^4) = (c_{\Delta-\Gamma}{}^4)^{-1}$, it follows that $\mathrm{ad}(c_{\Delta-\Gamma})^4$ preserves $\mathfrak{p}^C$ and that $\mathfrak{p}^C = \mathfrak{e} \oplus \mathfrak{f}$.

We have $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)\cdot\mathfrak{p}_\Gamma{}^C = 0$ and $\mathfrak{p}_\Gamma{}^C \subset \mathfrak{e}$, by definition of $\mathfrak{g}_\Gamma{}^C$. An application of [7, Lemma 5.3] to $(\mathfrak{g}_{\Delta-\Gamma}{}^0, \mathfrak{k}_{\Delta-\Gamma})$ shows that $\mathfrak{p}_{\Delta-\Gamma,1}{}^C$ is spanned by $(\pm i)$-eigenvectors of $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$, that $\mathfrak{p}_{\Delta-\Gamma}{}^C \cap \mathfrak{f}$ is spanned by $(\pm i/2)$-eigenvectors, and that $\mathfrak{p}_{\Delta-\Gamma}{}^C = \mathfrak{p}_{\Delta-\Gamma,1}{}^C + (\mathfrak{f} \cap \mathfrak{p}_{\Delta-\Gamma}{}^C)$. Let $\mathfrak{d}$ be the complexification of the orthogonal complement of $\mathfrak{p}_\Gamma + \mathfrak{p}_{\Delta-\Gamma}$ in $\mathfrak{p}$; it remains only to show that $\mathfrak{d} \subset \mathfrak{f}$ and that $\mathfrak{d}$ is spanned by $(\pm i/2)$-eigenspaces of $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$. As $\mathfrak{e}$ (resp. $\mathfrak{f}$) is the intersection with $\mathfrak{p}^C$ of the sum of the odd (resp. even) dimensional irreducible representation spaces of

$$\mathrm{ad}(\{X_{\Delta-\Gamma}, Y_{\Delta-\Gamma}, Z_{\Delta-\Gamma}{}^0\}),$$

we need only prove $\mathfrak{d}$ to be spanned by $(\pm i/2)$-eigenspaces, and then $\mathfrak{d} \subset \mathfrak{f}$ will follow.

$\mathfrak{d}$ is spanned by root vectors $E_{\pm\beta}$, β positive noncompact; thus we need only prove that $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)\cdot E_\beta = \mp (i/2)E_\beta$ for every noncompact positive root β with $E_\beta \in \mathfrak{d}$. By definition of $Z_{\Delta-\Gamma}{}^0$ and $\mathfrak{d}$, this is equivalent to the proof that, for every noncompact positive root β which is orthogonal neither to Γ nor to $\Delta - \Gamma$, we have $\sum_{\alpha \in \Delta-\Gamma} \langle \alpha, \beta \rangle = \pm 1$. This has been proved by Harish-Chandra [4, Lemmas 13-16]. *Q.E.D.*

4.5. Let ν^0 be conjugation of $\mathfrak{g}^C$ over $\mathfrak{g}^0$. As $\mathfrak{p}^0$ is spanned by the $X_\beta{}^0 = E_\beta + E_{-\beta}$ and the $Y_\beta{}^0 = -i(E_\beta - E_{-\beta})$ for the noncompact roots β, ν^0 exchanges E_β and $E_{-\beta}$, so $(I + \nu^0)iE_{-\beta} = -Y_\beta{}^0$ and $(I + \nu^0)E_{-\beta} = X_\beta{}^0$.

Let ν be the conjugation of $\mathfrak{g}^C$ over $\mathfrak{g}$ and observe that $\langle U, V\rangle_\nu = -\langle U, \nu V\rangle$ is a positive definite hermitian form on $\mathfrak{g}^C$ where $\langle\ ,\ \rangle$ denotes the Killing form. Let $\|\ \ \|$ denote operator norm relative to $\langle\ ,\ \rangle_\nu$ for linear transformations of $\mathfrak{g}^C$.

The following result is included for completeness. It was proved by C. C. Moore [8, Lemma 4.5] in a somewhat different manner. The idea of using operator norms is due to R. Hermann.

4.6. LEMMA. *Let* $\mathfrak{p}^0$ *be given the complex structure defined by* $\mathrm{ad}\,(Z)$

and define $\psi\colon \mathfrak{p}^- \to \mathfrak{p}^0$ *by* $\psi(E) = \frac{1}{2}(E + \nu^0 E)$. *Then* ψ *is an isomorphism of complex vector spaces, and*

$$\psi(D) = \{U \in \mathfrak{p}^0 \colon \|\operatorname{ad}(U)\| < 1\}.$$

Proof. The first statement is clear. ψ is $\operatorname{ad}(K)$-equivariant because $K \subset \exp(\mathfrak{g})$, and $\psi(\mathfrak{a}^-) = \mathfrak{a}^0$. Thus we need only prove $E \in D$ if and only if $\|\operatorname{ad}\psi(E)\| < 1$ for every $E \in \mathfrak{a}^-$. As $\mathfrak{a}^-$ consists of all $E = \sum_{\alpha\in\Delta} b_\alpha E_{-\alpha}$ with b_α real, and $E \in D$ if and only if each $|b_\alpha| < 1$, we need only prove that $\|\operatorname{ad}(\frac{1}{2}\sum b_\alpha X_\alpha{}^0)\| < 1$ is equivalent to the condition that each $|b_\alpha| < 1$.

Each $\mathfrak{g}_\beta{}^0$, $\beta \in \Delta$, has a nonzero element W_β with $[X_\beta{}^0, W_\beta] = 2W_\beta$; now $\operatorname{ad}(\frac{1}{2}\sum b_\alpha X_\alpha{}^0)\cdot W_\beta = b_\beta W_\beta$. Thus $\|\operatorname{ad}(\frac{1}{2}\sum b_\alpha X_\alpha{}^0)\| < 1$ implies that each $|b_\alpha| < 1$.

Suppose that each $|b_\alpha| < 1$; we will see that $\|\operatorname{ad}(\frac{1}{2}\sum b_\alpha X_\alpha{}^0)\| < 1$. As $Y \in \mathfrak{g}$, $\operatorname{ad}(\exp(\pi/4)Y)$ preserves operator norm; that element sends each $X_\alpha{}^0$ to H_α as seen by calculating in $\mathfrak{g}_\alpha{}^C$, so we need only prove $\|\operatorname{ad}(\frac{1}{2}\sum b_\alpha H_\alpha)\| < 1$. In other words, we need $|\sum_{\alpha\in\Delta} b_\alpha\langle\alpha,\beta\rangle| < 2$ for every root β. This now follows from [4, Lemmas 13-16] which say that, if $\langle\alpha,\beta\rangle \neq 0$ for some $\alpha \in \Delta$, then either $\langle\alpha,\beta\rangle = \pm 1$ and $\langle\alpha',\beta\rangle \neq 0$ for at most one other $\alpha' \in \Delta$, or $\langle\alpha, b\rangle = \pm 2$ and $\langle\alpha',\beta\rangle = 0$ for $\alpha \neq \alpha' \in \Delta$. *Q.E.D.*

We can now take the main step toward relating M_Γ to the boundary components of D. Here $\mathfrak{p}^-$ is endowed with the positive definite hermitian form $\langle\ ,\ \rangle_\nu$.

4.7. LEMMA. *Let* $\Gamma \subsetneqq \Delta$, *and define* $\mathfrak{e}_\Gamma{}^R$ *and* $\mathfrak{e}_\Gamma{}^C$ *to be the respective real and complex hyperplanes*[4] *in* $\mathfrak{p}^-$ *in which* o^Γ *is the point nearest to the origin. Then*

4.7.1. $$\bar{D} \cap [o^\Gamma + \mathfrak{p}_\Gamma{}^-] = \bar{D} \cap \mathfrak{e}_\Gamma{}^C = \bar{D} \cap \mathfrak{e}_\Gamma{}^R;$$

this set is the closure of $\xi^{-1}c_{\Delta-\Gamma}M_\Gamma$ *in* ∂D *and is a union of boundary components of* D; *it is the union of all sets of the form*

4.7.2. $$\operatorname{ad}(k)[\xi^{-1}c_{\Delta-\Sigma}M_\Sigma],\ k \in K_\Gamma, \Sigma \subset \Gamma.$$

Proof. This proof is close to an argument of Moore. $\psi = \frac{1}{2}(I + \nu^0)$ is a unitary transformation of $\mathfrak{p}^-$ onto $\mathfrak{p}^0$, so $\psi(\mathfrak{e}_\Gamma{}^R)$ consists of all

$$V = -\tfrac{1}{2}\sum_{\alpha\in\Delta-\Gamma} Y_\alpha{}^0 + U$$

where U is real-orthogonal to the first summand. Now decompose U

[4] real (resp. complex) affine subspaces of real (resp. complex) codimension 1.

$= U_1 + U_2$ where U_1 is real-orthogonal to $Y_\alpha{}^0$ for every $\alpha \in \Delta - \Gamma$, and where $U_2 = \sum_{\alpha \in \Delta - \Gamma} u_\alpha Y_\alpha{}^0$. The condition on U_1 gives

$$U_1 = \sum a_\beta X_\beta{}^0 + b_\beta Y_\beta{}^0$$

where the sum runs over all positive noncompact roots β, and where $b_\beta = 0$ in case $\beta \in \Delta - \Gamma$. As each $Y_\alpha{}^0$ has the same length, the condition on U implies $\sum_{\alpha \in \Delta - \Gamma} u_\alpha = 0$. Finally, by Lemma 4.6, $V \in \psi(\bar{D})$ if and only if $\| \operatorname{ad}(V) \| \leqq 1$.

Let $V \in \psi(\bar{D})$. We define $W_\alpha = X_\alpha{}^0 - 2Z_\alpha{}^0$ and $W = \sum_{\alpha \in \Delta - \Gamma} W_\alpha$. Now $[-\frac{1}{2} Y_\alpha{}^0, W] = W_\alpha$. This gives

$$\operatorname{ad}(V) \cdot W = W + [U_1, W],$$

and $[U_1, W]$ is real-orthogonal to W by definition of W and by $b_\beta = 0$ for $\beta \in \Delta - \Gamma$. As $\| \operatorname{ad}(V) \| \leqq 1$, we must have $\| \operatorname{ad}(V) \| = 1$ and

$$0 = [U_1, W] = 2 \operatorname{ad}(Z_{\Delta-\Gamma}{}^0) \cdot U_1 + F$$

where $F \in \mathfrak{k}$. This yields $\operatorname{ad}(Z_{\Delta-\Gamma}{}^0) \cdot U_1 = 0$; now $U_1 \in \mathfrak{p}_\Gamma{}^0$ by Lemma 4.4. As $U_2 \in \mathfrak{p}_\Gamma{}^0$ by construction, this proves $U \in \mathfrak{p}_\Gamma{}^0$. Thus $V \in \psi(\mathfrak{o}^\Gamma + \mathfrak{p}_\Gamma{}^-)$. We have just proved $\bar{D} \cap \mathfrak{e}_\Gamma{}^R \subset \bar{D} \cap [\mathfrak{o}^\Gamma + \mathfrak{p}_\Gamma{}^-]$; therefore (4.7.1) follows immediately. Lemma 4.2 shows that this set in the image by $\xi^{-1} c_{\Delta-\Gamma} \xi$ of the closure of $\xi^{-1} M_\Gamma$ in $\mathfrak{p}_\Gamma{}^-$, and the set lies in ∂D by the observation $\| \operatorname{ad}(V) \| = 1$ above; it follows that the set is the closure of $\xi^{-1} c_{\Delta-\Gamma} M_\Gamma$ in ∂D.

Let b be the complex linear functional on $\mathfrak{p}^-$ such that $b(E) = 1$ is the equation of $\mathfrak{e}_\Gamma{}^C$, and notice from the above paragraph that $\mathfrak{e}_\Gamma{}^C$ does not meet D. As D is preserved by the rotations $e^{i\theta}$, we then have $|b(E)| \leqq 1$ for every $E \in \bar{D}$. Now let $\mu \colon U \to \mathfrak{p}^-$ be an analytic arc in ∂D such that $\mu(U)$ meets $\mathfrak{e}_\Gamma{}^C$. Then the holomorphic function $b \circ \mu$ on U is bounded by 1 and this bound is achieved; thus $b \circ \mu$ is constant by the maximum modulus principle; in other words, $\mu(U) \subset \mathfrak{e}_\Gamma{}^C$. This proves that the set (4.7.1) is a union of boundary components of D.

The last statement follows by application of Lemma 4.3 to ∂D_Γ, $D_\Gamma = \xi^{-1}(M_\Gamma)$, and by the observation that $c_{\Delta-\Sigma} = c_{\Delta-\Gamma} \cdot c_{\Delta-\Sigma}$ for every $\Sigma \subset \Gamma$.

Q.E.D.

4.8. THEOREM. *The boundary components of D in $\mathfrak{p}^-$ are just the sets*

$$\operatorname{ad}(k)[\xi^{-1} c_{\Delta-\Gamma} M_\Gamma], \; k \in K, \; \Gamma \subsetneqq \Delta.$$

The boundary components of M in M^* are the sets

$$k(c_{\Delta-\Gamma}(M_\Gamma)), \; k \in K, \; \Gamma \subsetneqq \Delta.$$

Proof. The two statements are equivalent because ξ is an $\mathrm{ad}(K)$-equivariant complex analytic homeomorphism of a neighborhood of $\bar{D}$ in $\mathfrak{p}^-$ onto a neighborhood of $\bar{M}$ in M^*, carrying D onto M.

Every boundary component of D is a union of sets $\mathrm{ad}(k)[\xi^{-1}c_{\Delta-\Sigma}M_\Sigma]$, $k \in K$, $\Sigma \subsetneq \Delta$, and every such set lies in a boundary component of D, by Lemma 4.3. Thus it suffices to prove, given an analytic arc $\mu: U \to \mathfrak{p}^-$ in ∂D such that $\mu(U)$ meets $\xi^{-1}c_{\Delta-\Gamma}M_\Gamma$, that $\mu(U) \subset \xi^{-1}c_{\Delta-\Gamma}M_\Gamma$. Lemma 4.7 says that the closure of $\xi^{-1}c_{\Delta-\Gamma}M_\Gamma$ is a union of boundary components of D, so $\mu(U)$ is contained in that closure. Now define $\beta = \xi^{-1} \cdot c_{\Delta-\Gamma}{}^{-1} \cdot \xi \cdot \mu$; then $\beta: U \to \mathfrak{p}_\Gamma{}^-$ is an analytic arc in $\bar{D}_\Gamma$ which meets $D_\Gamma = \xi^{-1}(M_\Gamma)$, and we wish to prove that $\beta(U) \subset D_\Gamma$.

Suppose that $\beta(U)$ contains a point E of ∂D_Γ. Applying Lemma 4.3 to D_Γ we see that E is contained in a set $\mathrm{ad}(k)[\xi^{-1}c_{\Gamma-\Sigma}M_\Sigma]$, $k \in K_\Gamma$, $\Sigma \subsetneq \Gamma$. Applying Lemma 4.7 to D_Γ, we obtain a complex linear functional b on $\mathfrak{p}_\Gamma{}^-$ whose restriction to $\bar{D}_\Gamma$ attains its maximum at E; b is the linear functional specifying $\mathrm{ad}(k)e_\Sigma{}^C$. Now $b \circ \beta$ is a holomorphic function on U which attains its maximum, so $b \circ \beta$ is constant by the maximum modulus principle, whence

$$\beta(U) \subset \{F \in \bar{D}_\Gamma : b(F) = b(E)\} \subset \partial D_\Gamma.$$

This contradicts the fact that $\beta(U)$ meets D_Γ. This shows $\beta(U) \subset D_\Gamma$, and the Theorem is proved. *Q.E.D.*

4.9. Corollary. *The boundary components of D in $\mathfrak{p}^-$ are bounded symmetric domains in Harish-Chandra embedding, where the ambient space is a complex affine subspace of $\mathfrak{p}^-$ and the domain is the interior of the intersection of the ambient space with $\bar{D}$. The boundary components of M in M^* are hermitian symmetric spaces of noncompact type in Borel embedding, where the ambient space is a complex totally geodesic submanifold of M^* and the noncompact space is the interior of the intersection of the ambient space with $\bar{M}$.*

The first statement is immediate from Theorem 4.8 and Lemmas 4.2 and 4.7; the second statement follows upon mapping by ξ and applying Lemma 4.2.

4.10. Corollary. *Let $D = D_1 \times \cdots \times D_r$ be the decomposition of D as a product of irreducible domains. Then the boundary components of D are just the sets $F = F_1 \times \cdots \times F_r \neq D$ with F_j either equal to D_j or a boundary component of D_j; each of the bounded symmetric domains F_j is irreducible. The analogous result holds for the boundary components of M.*

Proof. Let F be a boundary component of D; then without loss of generality we may assume $F=\xi^{-1}c_{\Delta-\Gamma}M_\Gamma$ with $\Gamma \subsetneqq \Delta$. Let $\mathfrak{g}^0=\mathfrak{g}_1{}^0\oplus\cdots\oplus\mathfrak{g}_r{}^0$ be the decomposition as a sum of simple ideals, ordered so that the analytic subgroup of G^0 for $\mathfrak{g}_j{}^0$ is the connected group of analytic automorphisms of D_j. Then $\Delta=\Delta_1\cup\cdots\cup\Delta_r$ (disjoint) where Δ_j is a maximal set of strongly orthogonal noncompact roots of $\mathfrak{g}_j{}^0$. Define

$$\Gamma_j=\Gamma\cap\Delta_j \text{ and } F_j=\xi^{-1}c_{\Delta_j-\Gamma_j}(M_j\cap M_\Gamma).$$

Then $F=F_1\times\cdots\times F_r$, $F_j=D_j$ if $\Gamma_j=\Delta_j$, and F_j is a boundary component of D_j if $\Gamma_j\neq\Delta_j$.

Let $F=F_1\times\cdots\times F_r\neq D$ where F_j is D_j or a boundary component of D_j. The last part of the proof of Theorem 4.8 consisted of showing that D_Γ is an analytic arc component of $\bar{D}_\Gamma$; thus F_j is an analytic arc component of $\bar{D}_j$, so F is an analytic arc component of $\bar{D}$. Now $F\subset\partial D$ by construction, so F is an analytic arc component of ∂D, i.e., a boundary component of D.

To prove F_j irreducible we may assume D irreducible and $F_j=\xi^{-1}c_{\Delta-\Gamma}M_\Gamma$ with $\Gamma\subset\Delta$, and we need only prove that the effective part of $\mathfrak{g}_\Gamma{}^0$ is simple. It suffices to prove that the effective part of $\mathfrak{g}_\Gamma$ is simple. For this, we define W_Δ to be the subgroup of the Weyl group of G relative to $\mathfrak{h}$ consisting of the elements which preserve Δ as a set, and we define W_Γ to be the subgroup of W_Δ consisting of the elements which fix every element of $\Delta-\Gamma$. A result of C. C. Moore [8, Theorem 2] says that W_Δ induces the full group of permutations of Δ; thus W_Γ is transitive on Γ. Let U be the centralizer of $\mathfrak{h}_{\Delta-\Gamma}{}^-$ in G. $\exp(\mathfrak{h}_{\Delta-\Gamma}{}^-)$ is a torus because it is closed in $\exp(\mathfrak{h})$, so U is the centralizer of a torus. Now U is connected, the Weyl group of U relative to $\mathfrak{h}$ contains W_Γ (by definition of U), and G_Γ is the semisimple part of U (by definition of $\mathfrak{g}_\Gamma$); it follows that the Weyl group of G_Γ relative to $\mathfrak{h}\cap\mathfrak{g}_\Gamma$ is transitive on Γ. This proves that $\mathfrak{g}_\Gamma$ is simple. *Q.E.D.*

4.11. COROLLARY. *If M is of tube type, then each of its boundary components is of tube type. If M is irreducible and has a positive-dimensional boundary component of tube type, then M is of tube type.*

Proof. If M is of tube type, then the Cayley transform $c=c_\Delta$ has order 4. As $c=c_\Gamma\cdot c_{\Delta-\Gamma}$ and $\mathrm{ad}(c_{\Delta-\Gamma})|_{\mathfrak{g}_\Gamma}=1$, this implies $\mathrm{ad}(c_\Gamma)^4|_{\mathfrak{g}_\Gamma}=1$, and it follows that M_Γ is of tube type. Thus each boundary component is of tube type.

Before proving the second statement, we must check that M is of tube type whenever, for some $\Gamma\subset\Delta$, both M_Γ and $M_{\Delta-\Gamma}$ are of tube type. To see

this, we write $\mathfrak{p} = \mathfrak{p}_\Gamma + \mathfrak{p}_{\Delta-\Gamma} + \mathfrak{d}$ where $\mathfrak{d}$ is the orthogonal complement of $\mathfrak{p}_\Gamma + \mathfrak{p}_{\Delta-\Gamma}$. By Lemma 4.4, both $\mathrm{ad}(c_{\Delta-\Gamma})^4$ and $\mathrm{ad}(c_\Gamma)^4$ are -1 on $\mathfrak{d}$, so $\mathrm{ad}(c)^4 |_{\mathfrak{d}} = 1$. By hypothesis and the argument of the preceding paragraph, $\mathrm{ad}(c)$ is 1 on $\mathfrak{p}_\Gamma$ and on $\mathfrak{p}_{\Delta-\Gamma}$. Now $\mathrm{ad}(c)^4$ is 1 on $\mathfrak{p}$, and thus also on $\mathfrak{k} = [\mathfrak{p}, \mathfrak{p}]$, so $c^4 = 1$ and M is of tube type.

Let M be irreducible with a positive-dimensional boundary component of tube type. Then some M_Γ, $\phi \neq \Gamma \subsetneq \Delta$, is of tube type. Let $\alpha \in \Delta - \Gamma$ and $\beta \in \Gamma$, and define $\Phi = \Gamma \cup \{\alpha\}$ and $\Psi = \Phi - \{\beta\}$. A result of C. C. Moore [8, Theorem 2] shows that an element of the subgroup preserving $\mathfrak{h}^-$ in the Weyl group of $\mathfrak{g}^0$ send Γ to Ψ; thus M_Ψ is of tube type. Applying the first part of this Lemma to M_Ψ we see that $M_{\{\alpha\}}$ is of tube type. Applying the above paragraph to M_Φ with the decomposition $\Phi = \Gamma \cup \{\alpha\}$, now M_Φ is of tube type. Iterating the argument, $M_\Delta = M$ is seen to be of tube type.

Q.E.D.

4.12. COROLLARY. *For a bounded symmetric domain in Harish-Chandra embedding, a boundary component of a boundary component is a boundary component.*

This is immediate from Theorem 4.8 and from (4.7.2) in Lemma 4.7.

Lemma 4.4, Theorem 4.8 and Corollaries 4.11 and 4.12 allow us to list the boundary components. Here we say that two boundary components are of the *same type* if an element of G^0 sends one to the other.

4.13. THEOREM. *Let D be an irreducible bounded symmetric domain of rank m in Harish-Chandra embedding. For each integer r, $0 \leqq r < m$, there is just one type D_r of boundary component of D which has rank r as a symmetric space. D_0 is a single point and the other D_r are given as follows.*

4.13.1. $D = SU^m(2m+k)/S(U(m) \times U(m+k))$, $k \geqq 0$. *Then* $D_r = SU^r(2r+k)/S(U(r) \times U(r+k))$.

4.13.2. $D = SO^*(4m)/U(2m)$. *Then* $D_r = SO^*(4r)/U(2r)$.

4.13.3. $D = SO^*(4m+2)/U(2m+1)$.
Then $D_r = SO^*(4r+2)/U(2r+1)$.

4.13.4. $D = Sp(m,R)/U(m)$. *Then* $D_r = Sp(r,R)/U(r)$.

4.13.5. $D = SO^2(n+2)/SO(2) \times SO(n)$, $n > 2$; *here* $m = 2$. D_1 *is the unit disc in* C^1.

4.13.6. $D = E_6/SO(10) \cdot SO(2)$; *here* $m = 2$. M_1 *is the open unit ball in* C^5.

4.13.7. $D = E_7/E_6 \cdot SO(2)$; *here* $m = 3$. D_1 *is the unit disc in* C^1, *and* $D_2 = SO^2(12)/SO(2) \times SO(10)$.

Remark. For the classical domains the statement is due to L. K. Hua and K. H. Look [6], and the proof is due to Satake [10]. The result is new for the exceptional domains.

Remark. There are some duplications. For example $SO^*(8)/U(4) = SO^2(8)/SO(2) \times SO(6)$ and $SO^*(4)/U(2)$ is the unit disc in C^1.

Remark. The classification is not necessary for the first assertion. That assertion follows from Theorem 4.8 and transitivity of the small Weyl group on the collection of all subsets of r elements in $\mathrm{ad}(c)^2\Delta$.

Proof. The domains D listed exhaust the class of irreducible noncompact non-Euclidean hermitian symmetric spaces, according to É. Cartan. Here the domains of tube type are (4.13.1) for $k = 0$, (4.13.2), (4.13.4), (4.13.5) and (4.13.7). Now assertion (4.13.5) is immediate from Corollary 4.11 because the unit disc is the only tube-type domain of rank 1.

Let $\alpha \in \Delta$ and define $\Gamma = \Delta - \{\alpha\}$. The fixed point set of $\mathrm{ad}(c_\alpha)^4$ on $\mathfrak{g}_{\{\alpha\}}$ is of the form $\mathfrak{g}_{\{\alpha\},1} \oplus \mathfrak{l}_{\{\alpha\},2}$ where the second summand is in $\mathfrak{k}$ and the first is equal to $[\mathfrak{p}_{\{\alpha\},1}, \mathfrak{p}_{\{\alpha\},1}] + \mathfrak{p}_{\{\alpha\},1}$. Part 4 of [7, Theorem 4.9] shows that $\mathfrak{g}_{\{\alpha\},1}$ is of Cartan classification type $\mathfrak{a}_1$. Now Lemma 4.4 gives a direct sum decomposition $\mathfrak{p} = \mathfrak{p}_{\{\alpha\},1} + \mathfrak{p}_1 + \mathfrak{f}$ where $\mathrm{ad}(c_\alpha)^4$ is $+1$ on the first two summands and -1 on the third. Consider the decomposition $\mathfrak{g} = \mathfrak{u} + \mathfrak{v}$ into $+1$ and -1 eigenspaces of $\mathrm{ad}(c_\alpha)^4$; it follows that $\mathfrak{u} = \mathfrak{a}_1 \oplus \mathfrak{g}_1 \oplus \mathfrak{w}$ (direct sum of ideals) with $\mathfrak{w} \subset \mathfrak{k}$. As $\mathrm{ad}(c_\alpha)^4$ is an inner automorphism of $\mathfrak{g}$, we have proved: $\mathfrak{g}$ *has a symmetric subalgebra* $\mathfrak{u}$ *of maximal rank which has* $\mathfrak{a}_1$ *and* $\mathfrak{g}_\Gamma$ *as distinct simple ideals.*

Let $D = SU^m(2m+k)/S(U(m) \times U(m+k))$. Then $\mathfrak{g} = \mathfrak{a}_{2m+k-1}$, so the only possibility is $\mathfrak{u} = \mathfrak{a}_1 \oplus \mathfrak{a}_{2m+k-3} \oplus$ (1-dimensional abelian). Thus $\mathfrak{g}_\Gamma = \mathfrak{a}_{2(m-1)+k-1}$. As D_Γ has rank $m-1$, (4.13.1) follows.

Let $D = SO^*(2n)/U(n)$ where $n = 2m$ or $2m+1$. Then $\mathfrak{g} = \mathfrak{d}_n$ and $\mathfrak{u} = \mathfrak{d}_2 \oplus \mathfrak{d}_{n-2}$ is the only possibility; here observe that $\mathfrak{d}_2 = \mathfrak{a}_1 \oplus \mathfrak{a}_1$. If $\mathfrak{g}_\Gamma = \mathfrak{a}_1$, then $m-1 = 1$, and $n = 2m$ by Corollary 4.11, so $n = 4$ and $D_1 =$ (unit disc) $= SO^*(4)/U(2)$; conversely, if $n = 4$, then $\mathfrak{d}_{n-2} = \mathfrak{a}_1 \oplus \mathfrak{a}_1$ so $\mathfrak{g}_\Gamma = \mathfrak{d}_{n-2}$. We must check that

$$D_\Gamma = SO^*(2[n-2])/U(n-2).$$

If $m-1 > 2$ then this is true because the rank of D_Γ is too large to allow D_Γ to be of type (4.13.5); it is true if $m-1 = 2$ and $n = 2m+1$, for then

D_Γ cannot be of type (4.13.5) by Corollary 4.11; it is true for $m-1=2$ and $n=2m$ because $SO^*(8)/U(4)=SO^2(8)/SO(2)\times SO(6)$. Now (4.13.2) and (4.13.3) are proved.

Let $D=Sp(m,R)/U(m)$. Then $\mathfrak{g}=\mathfrak{c}_m$ so $\mathfrak{u}=\mathfrak{c}_1\oplus\mathfrak{c}_{m-1}$ and $\mathfrak{g}_\Gamma=\mathfrak{c}_{m-1}$. Thus $D_\Gamma=Sp(m-1,R)/U(m-1)$. This proves (4.13.4).

Let $D=E_6/SO(10)\cdot SO(2)$. Then $\mathfrak{g}=\mathfrak{e}_6$ so $\mathfrak{u}=\mathfrak{a}_1\oplus\mathfrak{a}_5$; thus $\mathfrak{g}_\Gamma=\mathfrak{a}_5$. As D_Γ is of rank 1, (4.13.6) is proved.

Let $D=E_7/E_6\cdot SO(2)$. Then $\mathfrak{g}=\mathfrak{e}_7$ so $\mathfrak{u}=\mathfrak{a}_1\oplus\mathfrak{d}_6$; thus $\mathfrak{g}_\Gamma=\mathfrak{d}_6$. As D_Γ is of tube type, or because $SO^*(12)/U(6)$ has rank 3, now D_Γ must be $SO^2(12)/SO(2)\times SO(10)$. This proves the statement on D_2; the statement on D_1 follows either from Corollary 4.11 or from Corollary 4.12.

Q.E.D.

The following result shows how the boundary components are related to the limit points of geodesic rays. We work in M and M^* for convenience, but the result translates immediately to D and $\mathfrak{p}^-$.

4.14. THEOREM. *Given $y\in M$ and a boundary component F of M in M^*, there is a unique point $f\in F$ such that some geodesic ray of M from y tends to f.*

Proof. Let $U=\{k(x^\Gamma): k\in K,\ \Gamma\subsetneq\Delta\}$, and define V to be the set of all limit points in M^* of geodesic rays of M with initial point x. If $X'=\sum_{\alpha\in\Delta}t_\alpha X_\alpha\in\mathfrak{a}^0$, then the geodesic ray $\{\exp(sX')\cdot x\}_{s\geq 0}$ is given by $\exp(sX')\cdot x=\xi(\sum_{\alpha\in\Delta}\tanh(t_\alpha s)\cdot E_{-\alpha})$. Thus the limit point of the geodesic ray is $\xi(\sum_{\alpha\in\Delta}\epsilon_\alpha E_{-\alpha})$ where ϵ_α is 0, 1 or -1 as t_α is 0, positive or negative. We can find $k\in K$ such that $\mathrm{ad}(k)X'=\sum|t_\alpha|\cdot X_\alpha$; now the limit point is $\mathrm{ad}(k)^{-1}\cdot x^\Gamma$ where $\Gamma=\{\alpha\in\Delta: t_\alpha=0\}$. This proves $U=V$.

We have $g\in G^0$ with $g(y)=x$, and $k\in K$ with $k(gF)=c_{\Delta-\Gamma}M_\Gamma$ for some $\Gamma\subsetneq\Delta$, so we may assume that $y=x$ and $F=c_{\Delta-\Gamma}M_\Gamma$. Now we need only prove that $c_{\Delta-\Gamma}M_\Gamma\cap\{k(x^\Sigma)$: x^Γ is the only element of $k\in K,\Sigma\subsetneq\Delta\}$. If $k(x^\Sigma)\in c_{\Delta-\Gamma}M_\Gamma$, then $c_{\Delta-\Gamma}M_\Gamma$ must coincide with $kc_{\Delta-\Sigma}M_\Sigma$, for both are boundary components containing $k(x^\Sigma)$. Lemma 4.7 shows that $o^\Gamma=\xi^{-1}(x^\Gamma)$ is closer to the origin of $\mathfrak{p}^-$ than any other point of $\xi^{-1}c_{\Delta-\Gamma}M_\Gamma$, and $\xi^{-1}(kx^\Sigma)$ is closer to the origin of $\mathfrak{p}^-$ than any other point of $\xi^{-1}kc_{\Delta-\Sigma}M_\Sigma$. Thus $k(x^\Sigma)=x^\Gamma$. *Q.E.D.*

5. The space of boundary components of a given type.

5.1. We say that two boundary components of D (or M) are of the *same type* if an element of G^0 carries one to the other, and we say that a

boundary component is of *type* Γ ($\Gamma \subsetneq \Delta$) if it is of the same type as $(\xi^{-1}c_{\Delta-\Gamma}\xi)D_\Gamma$ (or $c_{\Delta-\Gamma}M_\Gamma$). Here we remark

LEMMA. *Let* $\Gamma \subsetneq \Delta$ *and* $\Sigma \subsetneq \Delta$, *and suppose that* F *is a boundary component of type* Γ. *Then the following statements are equivalent.*

(i) F *is of type* Σ

(ii) $\mathrm{ad}(c)^2\Sigma$ *is equivalent to* $\mathrm{ad}(c)^2\Gamma$ *under the small Weyl group of* $(\mathfrak{g}^0, \mathfrak{k})$ *rel.* $\mathfrak{a}^0$

(iii) *For every simple ideal of* $\mathfrak{g}^0$, *both* Σ *and* Γ *contain the same number of roots of that ideal.*

Proof. Corollary 4.10 reduces the proof to the case where D is irreducible. Then (i) implies (iii) because Lemma 4.2 shows that the symmetric space rank of a component of type Γ is the number of elements of Γ. (iii) implies (ii) by [8, Theorem 2], and it is obvious that (ii) implies (i). *Q.E.D.*

Let $\Gamma \subsetneq \Delta$. We define S^Γ to be the set of all boundary components of M of type Γ, and we define $U^\Gamma \subset \partial M$ to be the union of all boundary components of type Γ. Similarly, $S_D{}^\Gamma$ denotes the set of boundary components of D of type Γ, and $U_D{}^\Gamma \subset \partial D$ is the union. These two notions coincide in the case where Γ is the empty set ϕ; there we have

$S^\emptyset = U^\emptyset = \check{S}$, Bergman-Šilov boundary of M in M^*

$S_D{}^\emptyset = U_D{}^\emptyset = \check{S}_D$, Bergman-Šilov boundary of D in $\mathfrak{p}^-$.

Theorem 4.8 and the Lemma above show that K acts transitively on S^Γ (resp. $S_D{}^\Gamma$). Let L^Γ denote the isotropy subgroup of K on $c_{\Delta-\Gamma}M_\Gamma \in S^\Gamma$ (resp. on $\xi^{-1}c_{\Delta-\Gamma}M_\Gamma \in S_D{}^\Gamma$; it is the same subgroup). Then L^Γ is the set of all elements of K which preserve the closure of $c_{\Delta-\Gamma}M_\Gamma$ in the compact set ∂M, and it follows that L^Γ is closed in K. Now we have identifications $S^\Gamma \cong K/L^\Gamma \cong S_D{}^\Gamma$, so S^Γ and $S_D{}^\Gamma$ are real analytic manifolds, homogeneous spaces of K.

As a final preliminary remark we observe that K cannot be transitive on U^Γ or $U_D{}^\Gamma$ for $\Gamma \neq \phi$, because any orbit of K is compact and Lemma 4.7 gives us the closures

$$\overline{U^\Gamma} = \bigcup_{\Sigma \subset \Gamma} U^\Sigma \text{ and } \overline{U_D{}^\Gamma} = \bigcup_{\Sigma \subset \Gamma} U_D{}^\Sigma.$$

5.2. LEMMA. *Let* $k \in K$. *If* k *preserves* $c_{\Delta-\Gamma}M_\Gamma$, *then* $k(x^\Gamma) = x^\Gamma$.

Proof. L^Γ is a compact group of isometries of $c_{\Delta-\Gamma}M_\Gamma$, so it has a

stationary point. But $K_\Gamma \subset L^\Gamma$, and x^Γ is the unique stationary point of K_Γ on $c_{\Delta-\Gamma}M_\Gamma$. This proves that x^Γ is stationary under L^Γ. *Q.E.D.*

5.3. Let $\pi\colon U^\Gamma \to S^\Gamma$ be the natural projection. This is a differentiable bundle with fibre M_Γ and group $G_\Gamma{}^0$, and K is transitive on the base. Lemma 5.2 may be paraphrased as: $kc_{\Delta-\Gamma}M_\Gamma \to k(x^\Gamma)$ *is a K-equivariant global section of the bundle* $U^\Gamma \to S^\Gamma$. Lemma 5.2 also allows us to identify S^Γ with $K(x^\Gamma)$.

5.4. *Definitions.* We will decompose $\mathfrak{g}^0$ under $c_{\Delta-\Gamma}$ in order to study S^Γ. Let $\tau_{\Delta-\Gamma} = \mathrm{ad}(c_{\Delta-\Gamma})^2$. We define:

$\mathfrak{g}^\Gamma$: the set of all elements of $\mathfrak{g}$ fixed under $\tau_{\Delta-\Gamma}{}^2$;
$\mathfrak{k}^\Gamma = \mathfrak{g}^\Gamma \cap \mathfrak{k}$;
$\mathfrak{p}_1{}^\Gamma = \mathfrak{g}^\Gamma \cap \mathfrak{p}$;
$\mathfrak{k}_1{}^\Gamma = [\mathfrak{p}_1{}^\Gamma, \mathfrak{p}_1{}^\Gamma]$;
$\mathfrak{g}_1{}^\Gamma = \mathfrak{k}_1{}^\Gamma + \mathfrak{p}_1{}^\Gamma$;
$\mathfrak{l}_2{}^\Gamma$: the centralizer of $\mathfrak{g}_1{}^\Gamma$ in $\mathfrak{g}^\Gamma$.

Here $\mathfrak{g}^\Gamma$ is a subalgebra of $\mathfrak{g}$, and $\mathfrak{g}^\Gamma = \mathfrak{k}^\Gamma + \mathfrak{p}_1{}^\Gamma$ because $\tau_{\Delta-\Gamma}{}^2$ preserves both $\mathfrak{k}$ and $\mathfrak{p}$. $\mathfrak{l}_2{}^\Gamma$ is the centralizer of $\mathfrak{p}_1{}^\Gamma$ in $\mathfrak{g}^\Gamma$, by the Jacobi identity; the decomposition theory of orthogonal involutive Lie algebras now implies that $\mathfrak{k}^\Gamma = \mathfrak{k}_1{}^\Gamma + \mathfrak{l}_2{}^\Gamma$ and $\mathfrak{g}^\Gamma = \mathfrak{g}_1{}^\Gamma \oplus \mathfrak{l}_2{}^\Gamma$, direct sums of ideals.

$\tau_{\Delta-\Gamma}$ preserves and has square I on $\mathfrak{k}_1{}^\Gamma$; its square preserves and has square I on $\mathfrak{k}$ and $\mathfrak{p}$. Thus we define:

$\mathfrak{l}_1{}^\Gamma$: the $(+1)$-eigenspace of $\tau_{\Delta-\Gamma}$ on $\mathfrak{k}_1{}^\Gamma$;
$\mathfrak{q}_1{}^\Gamma$: the (-1)-eigenspace of $\tau_{\Delta-\Gamma}$ on $\mathfrak{k}_1{}^\Gamma$;
$\mathfrak{q}_2{}^\Gamma$: the (-1)-eigenspace of $\tau_{\Delta-\Gamma}{}^2$ on $\mathfrak{k}$;
$\mathfrak{p}_2{}^\Gamma$: the (-1)-eigenspace of $\tau_{\Delta-\Gamma}{}^2$ on $\mathfrak{p}$;

$\mathfrak{l}^\Gamma = \mathfrak{l}_1{}^\Gamma + \mathfrak{l}_2{}^\Gamma$ and $\mathfrak{q}^\Gamma = \mathfrak{q}_1{}^\Gamma + \mathfrak{q}_2{}^\Gamma$.

Now we have $\mathfrak{k}_1{}^\Gamma = \mathfrak{l}_1{}^\Gamma + \mathfrak{q}_1{}^\Gamma$, $\mathfrak{k}^\Gamma = \mathfrak{l}^\Gamma + \mathfrak{q}_1{}^\Gamma$, $\mathfrak{k} = \mathfrak{l}^\Gamma + \mathfrak{q}^\Gamma$, and $\mathfrak{p} = \mathfrak{p}_1{}^\Gamma + \mathfrak{p}_2{}^\Gamma$.

We finally define some related subalgebras

$$\begin{aligned}\mathfrak{g}_1{}^{\Gamma,0} &= \mathfrak{k}_1{}^\Gamma + i\mathfrak{p}_1{}^\Gamma\\ \mathfrak{g}^{\Gamma,0} &= \mathfrak{k}^\Gamma + i\mathfrak{p}_1{}^\Gamma\\ \mathfrak{k}_1{}^{\Gamma *} &= \mathfrak{l}_1{}^\Gamma + i\mathfrak{q}_1{}^\Gamma\\ \mathfrak{k}^{\Gamma *} &= \mathfrak{l}^\Gamma + i\mathfrak{q}_1{}^\Gamma = \mathfrak{l}_2{}^\Gamma \oplus \mathfrak{k}_1{}^{\Gamma *}\end{aligned}$$

of $\mathfrak{g}^C$.

Latin letters denote the corresponding analytic subgroups of G^C, except that L^Γ was already defined to be the isotropy subgroup of K at x^Γ and

L_1^Γ will be the isotropy subgroup of K_1^Γ at x^Γ. We will justify this exception by checking that $\mathfrak{l}^\Gamma$ is the Lie algebra of L^Γ. As $L^\Gamma \subset G$ and $c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1}$ is the symmetry of $M^\#$ at x^Γ, this check is reduced to seeing that $\mathfrak{l}^\Gamma$ is the fixed point set of $\mathrm{ad}(c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1})$ on $\mathfrak{k}$. To prove the latter, we first observe that $\mathfrak{l}^\Gamma$ is the fixed point set of $\tau_{\Delta-\Gamma}$ on $\mathfrak{k}$, for the fixed point set is in $\mathfrak{l}^\Gamma$ by definition of $\mathfrak{q}^\Gamma$, the fixed point set contains $\mathfrak{l}_1^\Gamma$ by definition and the fixed point set contains $\mathfrak{l}_2^\Gamma$ as a consequence of $c_{\Delta-\Gamma} \in G_1^\Gamma$. Now let $V \in \mathfrak{k}$ and observe that

$$\begin{aligned}\mathrm{ad}(c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1}) \cdot V &= \mathrm{ad}(c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1}) \cdot \mathrm{ad}(s^{-1}) \cdot V \\ &= \mathrm{ad}(c_{\Delta-\Gamma}) \cdot \mathrm{ad}(\mathrm{ad}(s) c_{\Delta-\Gamma}) \cdot V = \tau_{\Delta-\Gamma}(V).\end{aligned}$$

Our assertion follows.

5.5. LEMMA. *$M^\Gamma = G^{\Gamma,0}(x)$ is a hermitian symmetric subspace of M. L_2^Γ is the identity component of the kernel of the action of $G^{\Gamma,0}$ on M^Γ, $G_1^{\Gamma,0}$ is (locally) the connected group of analytic automorphisms of M^Γ, and $\mathfrak{g}_1^\Gamma$ and $\mathfrak{g}_1^{\Gamma,0}$ are semisimple. The Cayley transform on M^Γ is $c = c_\Delta \in G_1^\Gamma$, M^Γ is of tube type if $\Gamma = \phi$, and M^Γ is of tube type if and only if M_Γ is of tube type when $\Gamma \neq \phi$.*

Proof. $\sum_{\alpha \in \Delta} \mathfrak{g}_\alpha \subset \mathfrak{g}^\Gamma$ by construction. Z is the sum of its projections Z' and Z^0 on $\mathfrak{h}^+$ and $\mathfrak{h}^-$, $\mathfrak{h}^- \subset \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha$, and $\mathfrak{h}^+$ is centralized by each c_α. This proves $Z \in \mathfrak{g}^\Gamma$, so $Z \in \mathfrak{g}^{\Gamma,0}$ and it follows that M^Γ is a sub hermitian symmetric space of M. The statement on L_2^Γ is immediate from the definition of $\mathfrak{l}_2^\Gamma$, and the assertions on $G_1^{\Gamma,0}$, $\mathfrak{g}_1^{\Gamma,0}$ and $\mathfrak{g}_1^\Gamma$ follow. The next statement follows from $\mathfrak{a}^0 \subset \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha^0 \subset \mathfrak{g}_1^{\Gamma,0}$, which is a consequence of $\sum_{\alpha \in \Delta} \mathfrak{g}_\alpha \subset \mathfrak{g}^\Gamma$ and the fact that $\mathfrak{g}_\alpha \cap \mathfrak{p}$ generates $\mathfrak{g}_\alpha$. The remaining assertions are immediate from Lemma 4.4. *Q.E.D.*

5.6. LEMMA. *$\tau_{\Delta-\Gamma}$ interchanges $\mathfrak{p}_2^\Gamma$ and $\mathfrak{q}_2^\Gamma$; $\mathrm{ad}(c_{\Delta-\Gamma})$ interchanges $\mathfrak{q}_1^\Gamma$ with the (-1)-eigenspace of $\tau_{\Delta-\Gamma}$ on $\mathfrak{p}_1^\Gamma$; $\mathfrak{p}_1^\Gamma = \mathfrak{p}_\Gamma + \mathfrak{p}_{\Delta-\Gamma,1}$ where $\tau_{\Delta-\Gamma}$ is $+1$ on $\mathfrak{p}_\Gamma$ and has square 1 on $\mathfrak{p}_{\Delta-\Gamma,1}$, and where $J = \mathrm{ad}\, Z$ interchanges the (± 1)-eigenspaces of $\tau_{\Delta-\Gamma}$ on $\mathfrak{p}_{\Delta-\Gamma,1}$.*

Proof. Let $V \in \mathfrak{p}_2^\Gamma$ and $R \in \mathfrak{q}_2^\Gamma$. Then

$$\sigma\tau_{\Delta-\Gamma} V = \tau_{\Delta-\Gamma}^{-1}\sigma V = -\tau_{\Delta-\Gamma}^{-1} V = -\tau_{\Delta-\Gamma}\tau_{\Delta-\Gamma}^2 V = \tau_{\Delta-\Gamma} V$$

and

$$\sigma\tau_{\Delta-\Gamma} R = \tau_{\Delta-\Gamma}^{-1}\sigma R = \tau_{\Delta-\Gamma}^{-1} R = \tau_{\Delta-\Gamma}\tau_{\Delta-\Gamma}^2 R = -\tau_{\Delta-\Gamma} R.$$

Thus $\tau_{\Delta-\Gamma}(\mathfrak{p}_2{}^\Gamma) \subset \mathfrak{k}$ and $\tau_{\Delta-\Gamma}(\mathfrak{q}_2{}^\Gamma) \subset \mathfrak{p}$. Now $\tau_{\Delta-\Gamma}$ comutes with its own square, and this implies $\tau_{\Delta-\Gamma}(\mathfrak{p}_2{}^\Gamma) \subset \mathfrak{q}_2{}^\Gamma$ and $\tau_{\Delta-\Gamma}(\mathfrak{q}_2{}^\Gamma) \subset \mathfrak{p}_2{}^\Gamma$. Equality follows from dimension considerations. This proves the interchange statement for $\tau_{\Delta-\Gamma}$. The proof of the interchange statement for $\mathrm{ad}(c_{\Delta-\Gamma})$ is similar.

Lemma 4.4 shows $\mathfrak{p} = \mathfrak{p}_\Gamma + \mathfrak{p}_{\Delta-\Gamma,1} + \mathfrak{p}_2{}^\Gamma$, direct sum; thus we need only check that $J = \mathrm{ad}(Z)$ interchanges the (± 1)-eigenspaces of $\mathfrak{p}_{\Delta-\Gamma,1}$. This follows from the fact that Lemma 4.2 allows us to apply [7, Lemma 4.7] to $M_{\Delta-\Gamma}$. *Q.E.D.*

5.7. THEOREM. *Let L^Γ and $L_1{}^\Gamma$ be the isotropy subgroups of K and $K_1{}^\Gamma$ at $x^\Gamma = c_{\Delta-\Gamma}(x)$. Then:*

5.7.1. $S^\Gamma = K(x^\Gamma) \cong K/L^\Gamma$ *and* $\dim. S^\Gamma = \dim. \mathfrak{p}_2{}^\Gamma + \frac{1}{2}\dim. \mathfrak{p}_{\Delta-\Gamma,1}$.

5.7.2. $U^\Gamma = G^0(x^\Gamma)$ *and* $\dim. U^\Gamma = \dim. \mathfrak{p}_2{}^\Gamma + \frac{1}{2}\dim. \mathfrak{p}_{\Delta-\Gamma,1} + \dim. \mathfrak{p}_\Gamma$.

5.7.3. *$K(c_{\Delta-\Gamma}{}^2(x))$ is a complex totally geodesic submanifold of M^*, and is thus a compact herimitian symmetric space; K^Γ is the isotropy subgroup of K at $c_{\Delta-\Gamma}{}^2(x)$, so $K(c_{\Delta-\Gamma}{}^2(x)) \cong K/K^\Gamma$.*

5.7.4. *The map $k(x^\Gamma) \to k(c_{\Delta-\Gamma}{}^2(x))$ is a fibering of S^Γ over $K(c_{\Delta-\Gamma}{}^2(x))$; the fibre over $k(c_{\Delta-\Gamma}{}^2(x))$ is $kK_1{}^\Gamma(x^\Gamma)$, which is totally geodesic in M^*, Riemannian symmetric and isometric to $K_1{}^\Gamma/L_1{}^\Gamma$.*

5.7.5. *The following statements are equivalent:*

(i) *The partial Cayley transform $c_{\Delta-\Gamma}$ has order* 4, i.e., $\mathfrak{g} = \mathfrak{g}^\Gamma$, *i.e., $K(c_{\Delta-\Gamma}{}^2(x))$ is a single point.*

(ii) *S^Γ is a totally geodesic submanifold of M^* (in which case it is Riemannian symmetric and K induces the largest connected group of isometries).*

(iii) *Let $M = M_1 \times \cdots \times M_r$ be the decomposition into irreducible factors, and let $c_{\Delta-\Gamma}M_\Gamma = F_1 \times \cdots \times F_r$ be the corresponding decomposition of the boundary component $c_{\Delta-\Gamma}M_\Gamma$. Then for each j, either $F_j = M_j$, or M_j is of tube type and F_j is a point on its Bergman-Šilov boundary.*

Proof. $S^\Gamma = K(x^\Gamma)$ was observed in § 5.3, and $K(x^\Gamma) \cong K/L^\Gamma$ by definition of L^Γ. Now $\dim. S^\Gamma = \dim. K - \dim. L^\Gamma = \dim. \mathfrak{k} - \dim. \mathfrak{l}^\Gamma = \dim. \mathfrak{q}^\Gamma - \dim. \mathfrak{q}_1{}^\Gamma + \dim. \mathfrak{q}_2{}^\Gamma$. Lemma 5.6 shows that $\dim. \mathfrak{q}_2{}^\Gamma = \dim. \mathfrak{p}_2{}^\Gamma$ and $\dim. \mathfrak{q}_1{}^\Gamma = \frac{1}{2}\dim. \mathfrak{p}_{\Delta-\Gamma,1}$. This proves (5.7.1).

$U^\Gamma = G^0(M_\Gamma) = G^0(G_\Gamma(x^\Gamma)) = G^0(x^\Gamma)$, and Lemma 5.2 shows that $\dim. U^\Gamma - \dim. S^\Gamma = \dim. M_\Gamma = \dim. \mathfrak{p}_\Gamma$. Now (5.7.2) follows from (5.7.1).

The isotropy subalgebra of $\mathfrak{k}$ at $c_{\Delta-\Gamma}{}^2(x)$ is the fixed piont set in $\mathfrak{k}$ of

conjugation by the symmetry $c_{\Delta-\Gamma}^2 s c_{\Delta-\Gamma}^{-2} = c_{\Delta-\Gamma}^4 s$ there; thus $\mathfrak{k}^\Gamma$ is the isotropy subalgebra of $\mathfrak{k}$ at $c_{\Delta-\Gamma}^2(x)$. On the other hand, $\mathfrak{k} = \mathfrak{k}^\Gamma + \mathfrak{q}_2^\Gamma$, and conjugation by the symmetry is -1 on $\mathfrak{q}_2^\Gamma$; thus the orbit $K(c_{\Delta-\Gamma}^2(x))$ is totally geodesic in M^*. The complex structure operator at $c_{\Delta-\Gamma}^2(x)$ is $\tau_{\Delta-\Gamma}(Z)$. $[\tau_{\Delta-\Gamma}Z, \mathfrak{q}_2^\Gamma] = \tau_{\Delta-\Gamma}^{-1}[Z, \tau_{\Delta-\Gamma}\mathfrak{q}_2^\Gamma] = \tau_{\Delta-\Gamma}^{-1}[Z, \mathfrak{p}_2^\Gamma] = \tau_{\Delta-\Gamma}^{-1}(\mathfrak{p}_2^\Gamma) = \mathfrak{q}_2^\Gamma$. Thus $K(c_{\Delta-\Gamma}^2(x))$ is a complex submanifold of M^*. $K(c_{\Delta-\Gamma}^2(x))$ is simply connected; this is seen in the irreducible case because a local toral factor would be a coset space of the (one real dimensional) center of K, and the assertion follows in general. Now the isotropy subgroup of K at $c_{\Delta-\Gamma}^2(x)$ is connected; as its Lie algebra is $\mathfrak{k}^\Gamma$, it must be the analytic group K^Γ. This proves (5.7.3).

The map $S^\Gamma \to K(c_{\Delta-\Gamma}^2(x))$ is given by the map $kL^\Gamma \to kK^\Gamma$ of K/L^Γ onto K/K^Γ; to prove it to be well defined, we must check that $L^\Gamma \subset K^\Gamma$ (although we do not yet know that L^Γ is connected). K^Γ is the identity component of V, where V is the full centralizer of $\tau_{\Delta-\Gamma}^2$ in K. As K/V is hermitian symmetric without locally euclidean factor, as checked in the paragraph above, it is simply connected. Thus V is connected, and now $K^\Gamma = V$. On the other hand, $L^\Gamma = K \cap \mathrm{ad}(c_{\Delta-\Gamma})K$ is contained in the centralizer of $\tau_{\Delta-\Gamma}$ in K. Thus $L^\Gamma \subset K^\Gamma$. Now $S^\Gamma \to K(c_{\Delta-\Gamma}^2(x))$ is a well-defined fibering. The fibre over $k(c_{\Delta-\Gamma}^2(x))$ is $kK^\Gamma(x^\Gamma) = k\cdot K_1^\Gamma \cdot L_2^\Gamma(x^\Gamma) = kK_1^\Gamma(x^\Gamma)$. $K_1^\Gamma(x^\Gamma)$ is totally geodesic in M^*, because $c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1}$ is the symmetry at x^Γ, and because $\mathrm{ad}(c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1})K_1^\Gamma = \tau_{\Delta-\Gamma}K_1^\Gamma = K_1^\Gamma$. Now kK_1^Γ is totally geodesic in M^*. We have proved (5.7.4).

Let $c_{\Delta-\Gamma}^4 = 1$. Then $\mathfrak{g} = \mathfrak{g}^\Gamma$, so in particular $\mathfrak{k} = \mathfrak{k}^\Gamma$ and $K(c_{\Delta-\Gamma}^2(x)) \cong K/K^\Gamma$ is a single point. If $\mathfrak{k} = \mathfrak{k}^\Gamma$, then $\mathfrak{q}_2^\Gamma = 0$, so $\mathfrak{p}_2^\Gamma = 0$ by Lemma 5.6, whence $\mathfrak{g} = \mathfrak{g}^\Gamma$ and $c_{\Delta-\Gamma}^4 = 1$. Now the conditions of (i) of (5.7.5) are equivalent.

Assume (i). Then s commutes with $c_{\Delta-\Gamma}^2$ because $c_{\Delta-\Gamma}^4 = 1$, so $\tau_{\Delta-\Gamma}(\mathfrak{k}) = \mathfrak{k}$. As $\tau_{\Delta-\Gamma}$ coincides with $\mathrm{ad}(c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1})$ on $\mathfrak{k}$, $S^\Gamma = K(x^\Gamma)$ is totally geodesic in M^*, which is (ii). Assume (ii). If M is irreducible then K is the largest connected subgroup of G which preserves $K(x^\Gamma)$, by maximality of $\mathfrak{k}$ in $\mathfrak{g}$; now $\mathfrak{k} = \mathrm{ad}(c_{\Delta-\Gamma} s c_{\Delta-\Gamma}^{-1})\mathfrak{k} = \tau_{\Delta-\Gamma}(\mathfrak{k})$ by (ii), and (i) follows via Lemma 5.2 from $\mathfrak{q}_2^\Gamma = 0 = \mathfrak{p}_2^\Gamma$. Now (i) is equivalent to (ii) in (5.7.5).

For the equivalence of (i) and (iii) we may assume M irreducible. Assume (iii), then M_Γ is a point and M is of tube type, so $c_{\Delta-\Gamma} = c$ and [7, Theorem 4.9] $c^4 = 1$, proving (i). Assume (i). Then $M = M^\Gamma$. As $M^\Gamma = M_\Gamma \times M_{\Delta-\Gamma,1}$ by $\mathfrak{p}_1^\Gamma = \mathfrak{p}_\Gamma + \mathfrak{p}_{\Delta-\Gamma,1}$, and as M is irreducible, we must

have $\Gamma=\phi$ or $\Gamma=\Delta$; (iii) follows. Now (i) and (iii) are equivalent in (5.7.5). *Q.E.D.*

5.8. COROLLARY. *The fundamental group $\pi_1(S^\Gamma)$ is the direct product of a finite abelian group and a group which is free abelian with one generator for each tube type irreducible factor of M whose Bergman-Šilov boundary is a direct factor of $c_{\Delta-\Gamma}M_\Gamma$. In particular the first Betti number of S^Γ is the number of irreducible tube type factors of M whose Bergman-Šilov boundary is a factor of $c_{\Delta-\Gamma}M_\Gamma$.*

Proof. We may assume M irreducible. Now $Z\in\mathfrak{q}_1{}^\Gamma$ if and only if M is of tube type and $\Gamma=\phi$, for $Z=Z'+Z_\Gamma{}^0+Z_{\Delta-\Gamma}{}^0$ where $Z'+Z_\Gamma{}^0\in\mathfrak{l}^\Gamma$, $Z_{\Delta-\Gamma}{}^0\in\mathfrak{q}_1{}^\Gamma$, and $Z'=0$ if and only if M is of tube type. Let $\mathfrak{k}_{ss}$ be the derived algebra of $\mathfrak{k}$. As $\dim.\mathfrak{k}-\dim.\mathfrak{k}_{ss}=1$, $\mathfrak{k}_{ss}\perp Z$ under the Killing form of $\mathfrak{g}$, and $\mathfrak{l}^\Gamma\perp\mathfrak{q}_1{}^\Gamma$, it follows that

(i) if M is of tube type and $\Gamma=\phi$ then $\mathfrak{l}^\Gamma\subset\mathfrak{k}_{ss}$, and

(ii) otherwise $\mathfrak{k}_{ss}+\mathfrak{l}^\Gamma=\mathfrak{k}$.

We also have

(iii) $\pi_1(S^\Gamma)$ is abelian

as in [7, Theorem 4.11] because S^Γ is fibered over a hermitian symmetric space of a semisimple group with symmetric fibre. Now our assertion follows from some homotopy sequences as in [7, Theorem 4.11]. *Q.E.D.*

6. The stability group of a boundary component. G^0 is transitive both on the set S^Γ of boundary components of type Γ and on the union U^Γ of these boundary components. Let B^Γ be the set of all elements of G^0 which preserve $c_{\Delta-\Gamma}M_\Gamma\in S^\Gamma$ and define E^Γ to be the isotropy subgroup of G^0 at $x^\Gamma=c_{\Delta-\Gamma}(x)$. Now

$$S^\Gamma\cong G^0/B^\Gamma \text{ and } U^\Gamma\cong G^0/E^\Gamma.$$

We will study S^Γ and U^Γ by examining B^Γ and E^Γ.

6.1. We have $L^\Gamma\subset E^\Gamma\subset B^\Gamma$ because L^Γ is the isotropy subgroup of K at x^Γ and $x^\Gamma\in c_{\Delta-\Gamma}M_\Gamma$, and $L^\Gamma=K\cap B^\Gamma$ by Lemma 5.2. K is transitive on S^Γ so $G^0=KB^\Gamma$; now $G^0=B^\Gamma\cdot K$ and B^Γ is transitive on M, $B^\Gamma/L^\Gamma=M$. M being connected and acyclic, it follows that L^Γ is a maximal compact subgroup of B^Γ.

E^Γ is in general not transitive on M. For if $E^\Gamma(x)=M$, then $\dim.E^\Gamma=\dim.B^\Gamma$ because $K\cap B^\Gamma=L^\Gamma=K\cap E^\Gamma$, whence $E^\Gamma=B^\Gamma$ because it meets every component. Then $c_{\Delta-\Gamma}M_\Gamma=x^\Gamma$ because $G_\Gamma{}^0\subset B^\Gamma$ and so $\Gamma=\phi$.

Now L^Γ is maximal compact both in B^Γ and E^Γ, and these groups are

generated by L^Γ and their respective identity components $B_0{}^\Gamma$ and $E_0{}^\Gamma$. This brings the study of B^Γ and E^Γ down to the study of their Lie algebras $\mathfrak{b}^\Gamma$ and $\mathfrak{e}^\Gamma$. We will need some definitions in order to calculate these Lie algebras.

(6.1.1) If $\mathfrak{u}$ is a subspace of $\mathfrak{g}$ or $\mathfrak{g}^0$ and $\mathfrak{u}^C$ is the sum of $\mathfrak{h}^C$-root spaces, then $\mathfrak{u}^+$ (resp. $\mathfrak{u}^-$) denotes the sum of the positive (resp. negative) root spaces in $\mathfrak{u}^C$. This defines $\mathfrak{p}_i{}^{\Gamma\pm}$, $\mathfrak{q}_2{}^{\Gamma\pm}$, $\mathfrak{p}_\Sigma{}^\pm$ and $\mathfrak{p}_{\Sigma,i}{}^\pm$ ($i=1,2;\Sigma=\Gamma,\Delta-\Gamma$).
(6.1.2) $\mathfrak{r}_2{}^{\Gamma\pm}=\mathfrak{q}_2{}^{\Gamma\pm}+\mathfrak{p}_2{}^{\Gamma\pm}$ and $\mathfrak{r}^{\Gamma\pm}=\mathfrak{p}_{\Delta-\Gamma,1}{}^\pm+\mathfrak{r}_2{}^{\Gamma\pm}$, complex subspaces of $\mathfrak{g}^C$.
(6.1.3) $\mathfrak{n}_2{}^{\Gamma\pm}=\mathfrak{r}_2{}^{\Gamma\pm}\cap \operatorname{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$, $\mathfrak{n}_1{}^{\Gamma\pm}=\mathfrak{p}_{\Delta-\Gamma,1}{}^\pm\cap \operatorname{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$ and

$$\mathfrak{n}^{\Gamma\pm}=\mathfrak{n}_1{}^{\Gamma\pm}+\mathfrak{n}_2{}^{\Gamma\pm}, \text{ real subspaces of } \mathfrak{g}^C.$$

(6.1.4) Recall $\mathfrak{k}_1{}^{\Gamma*}=\mathfrak{l}_1{}^\Gamma+i\mathfrak{q}_1{}^\Gamma$ and $\mathfrak{k}^{\Gamma*}=\mathfrak{l}^\Gamma+i\mathfrak{q}_1{}^\Gamma$.

Convention. From now on we assume that $\alpha>\beta$ for $\alpha\in\Gamma$ and $\beta\in\Delta-\Gamma$. This causes no loss of generality because [8, Theorem 2] on each irreducible factor of M the small Weyl group induces all permutations on the strongly orthogonal roots.

6.2. LEMMA. *$\operatorname{ad}(Z_{\Delta-\Gamma}{}^0)$ coincides with $\frac{1}{2}\operatorname{ad}(Z)$ on $\mathfrak{p}_2{}^{\Gamma C}$ and $\tau_{\Delta-\Gamma}$ interchanges $\mathfrak{p}_2{}^{\Gamma\pm}$ with $\mathfrak{q}_2{}^{\Gamma\pm}$.*

Proof. $Z=(Z'+Z_\Gamma{}^0)+Z_{\Delta-\Gamma}{}^0$ where $\tau_{\Delta-\Gamma}$ is $+1$ on the first summand and -1 on the second. Now $\tau_{\Delta-\Gamma}Z=Z-2Z_{\Delta-\Gamma}{}^0$. Let $E\in\mathfrak{p}_2{}^{\Gamma C}$; $E=\tau_{\Delta-\Gamma}Q$ with $Q\in\mathfrak{q}_2{}^{\Gamma C}$ by Lemma 5.6, and

$$(\operatorname{ad}(Z)-2\operatorname{ad}(Z_{\Delta-\Gamma}{}^0))E=\operatorname{ad}(\tau_{\Delta-\Gamma}Z)(\tau_{\Delta-\Gamma}Q)=\tau_{\Delta-\Gamma}\operatorname{ad}(Z)Q=0.$$

This proves the first statement.

We may now assume $\mathfrak{g}^0$ simple. Let $\Delta=\{\delta_1,\cdots,\delta_r\}$ with $\delta_1<\delta_2<\cdots<\delta_r$, so $\Gamma=\{\delta_{i+1},\cdots,\delta_r\}$ by hypothesis on the ordering of roots. It is known [8, Theorem 1] that the compact simple roots have restrictions 0, $\frac{1}{2}(\delta_2-\delta_1),\cdots,\frac{1}{2}(\delta_r-\delta_{r-1})$, and perhaps also $-\frac{1}{2}\delta_r$, to $\mathfrak{h}^-$. Thus $\operatorname{ad}(Z_{\Delta-\Gamma}{}^0)\cdot E_\beta$ is 0 or $(i/2)E_\beta$. It follows that $\operatorname{ad}(Z_{\Delta-\Gamma}{}^0)\cdot E_\gamma=ia_\gamma E_\gamma$ with $a_\gamma\geqq 0$ for every compact positive root γ. The first statement says that $\operatorname{ad}(Z_{\Delta-\Gamma}{}^0)$ is $\mp i/2$ on $\mathfrak{p}_2{}^{\Gamma\pm}$. As $\tau_{\Delta-\Gamma}{}^0(Z_{\Delta-\Gamma}{}^0)=-Z_{\Delta-\Gamma}{}^0$, it follows that $\operatorname{ad}(Z_{\Delta-\Gamma}{}^0)$ is $\pm i/2$ on $\tau_{\Delta-\Gamma}(\mathfrak{p}_2{}^{\Gamma\pm})$. Thus $\tau_{\Delta-\Gamma}(\mathfrak{p}_2{}^{\Gamma\pm})\subset\mathfrak{q}_2{}^{\Gamma\pm}$ and the interchange statement follows. *Q.E.D.*

6.3. LEMMA. *The eigenvalues and eigenspaces of* $\operatorname{ad}(-Y_{\Delta-\Gamma}{}^0)$ *are*:

eigenvalue	*eigenspace on* $\mathfrak{g}^C$	*eigenspace on* $\mathfrak{g}^0$
0	$\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma C}+\mathfrak{p}_\Gamma{}^C$	$\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma*}+\mathfrak{p}_\Gamma{}^0$
± 1	$\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{r}_2{}^{\Gamma\pm}$	$\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{n}_2{}^{\Gamma\pm}$
± 2	$\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{p}_{\Delta-\Gamma,1}{}^\pm$	$\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{n}_1{}^{\Gamma\pm}$

In particular, $\mathfrak{k}^{\Gamma *}=\mathfrak{l}_2{}^{\Gamma}\oplus\mathfrak{k}_1{}^{\Gamma *}=\mathfrak{k}^{\Gamma C}\cap\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$ *and is a real form of* $\mathfrak{k}^{\Gamma C}$, $\mathfrak{n}_2{}^{\Gamma\pm}$ *is a real form of* $\mathfrak{r}_2{}^{\Gamma\pm}$, *and* $\mathfrak{n}_1{}^{\Gamma\pm}$ *is a real form of* $\mathfrak{p}_{\Delta-\Gamma,1}{}^{\pm}$.

Proof. $\mathfrak{g}^C=\mathfrak{k}^{\Gamma C}+\mathfrak{q}_2{}^{\Gamma C}+\mathfrak{p}_2{}^{\Gamma C}+\mathfrak{p}_\Gamma{}^C+\mathfrak{p}_{\Delta-\Gamma,1}{}^C$. $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$ is $\mp i$ on $\mathfrak{p}_{\Delta-\Gamma,1}{}^{\pm}$ by Lemma 4.4, $\pm i/2$ on $\mathfrak{q}_2{}^{\Gamma\pm}+\mathfrak{p}_2{}^{\Gamma\mp}$ by Lemma 6.2, and 0 on $\mathfrak{p}_\Gamma{}^C$ by definition of $\mathfrak{g}_\Gamma{}^C$. $\mathfrak{k}^{\Gamma C}=\mathfrak{l}_2{}^{\Gamma C}+(\mathfrak{k}_\Gamma{}^C+\mathfrak{k}_{\Delta-\Gamma,1}{}^C)$ for $\mathfrak{p}^{\Gamma C}=\mathfrak{p}_\Gamma{}^C+\mathfrak{p}_{\Delta-\Gamma,1}{}^C$ by Lemma 4.4. Now $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)\mathfrak{k}^{\Gamma C}=0$ because $Z_{\Delta-\Gamma}{}^0$ is central in $\mathfrak{k}_{\Delta-\Gamma,1}{}^{0C}$, centralizes $\mathfrak{k}_\Gamma{}^C$ by construction, and centralizes $\mathfrak{l}_2{}^{\Gamma C}$ by $[\mathfrak{l}_2{}^\Gamma,\mathfrak{k}_1{}^\Gamma]=0$ and $\mathfrak{k}_{\Delta-\Gamma,1}\subset\mathfrak{k}_1{}^\Gamma$. Now we know the eigenvalues and eigenspaces of $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$ on $\mathfrak{g}^C$, and the assertions for $\mathrm{ad}(-Y_{\Delta-\Gamma}{}^0)$ follow from

$$\mathrm{ad}(c_{\Delta-\Gamma})^{-1}(2iZ_{\Delta-\Gamma}{}^0)=-Y_{\Delta-\Gamma}{}^0.$$

As $-Y_{\Delta-\Gamma}{}^0\in\mathfrak{g}^0$ and $\mathrm{ad}(-Y_{\Delta-\Gamma}{}^0)$ is a semisimple linear transformation with all eigenvalues real, every eigenspace of $\mathrm{ad}(-Y_{\Delta-\Gamma}{}^0)$ on $\mathfrak{g}^C$ is the complexification of its intersection with $\mathfrak{g}^0$. Thus we need only prove that

$$\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{p}_{\Delta-\Gamma,1}{}^{\pm}\cap\mathfrak{g}^0=\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{n}_1{}^{\Gamma\pm},$$

that

$$\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{r}_2{}^{\Gamma\pm}\cap\mathfrak{g}^0=\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{n}_2{}^{\Gamma\pm},$$

and that

$$(\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma C}+\mathfrak{p}_\Gamma{}^C)\cap\mathfrak{g}^0=\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma *}+\mathfrak{p}_\Gamma{}^0.$$

The first two equalities are immediate from the definitions of the $\mathfrak{n}_i{}^{\Gamma\pm}$, and $\mathfrak{p}_\Gamma{}^C\cap\mathfrak{g}^0=\mathfrak{p}_\Gamma{}^0$ by construction. Thus we need only prove that

$$\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma C}\cap\mathfrak{g}^0=\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma *}.$$

As $\mathfrak{k}^{\Gamma *}$ is a real form of $\mathfrak{k}^{\Gamma C}$, it suffices to check that $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma *}\subset\mathfrak{g}^0$.

$\mathfrak{k}^{\Gamma *}=\mathfrak{l}^\Gamma+i\mathfrak{q}_1{}^\Gamma$ and $\mathrm{ad}(c_{\Delta-\Gamma})$ is trivial on $\mathfrak{l}^\Gamma$. Thus $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{l}^\Gamma=\mathfrak{l}^\Gamma\subset\mathfrak{k}\subset\mathfrak{g}^0$. Lemma 5.6 says that $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}(i\mathfrak{q}_1{}^\Gamma)\subset i\mathfrak{p}_1{}^\Gamma\subset\mathfrak{p}^0\subset\mathfrak{g}^0$. Now $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma *}\subset\mathfrak{g}^0$ and the Lemma is proved. *Q.E.D.*

6.4. LEMMA. $[\mathfrak{q}_2{}^{\Gamma\mp},\mathfrak{p}_2{}^{\Gamma\pm}]\subset\mathfrak{p}_{\Delta-\Gamma,1}{}^{\pm}$, $\mathfrak{r}^{\Gamma\pm}$ *is a complex nilpotent subalgebra of degree 2 which is unipotent in the adjoint representation of* $\mathfrak{g}^C$, *and* $\mathfrak{n}^{\Gamma\pm}$ *is a real form of* $\mathfrak{r}^{\Gamma\pm}$.

Proof. $[\mathfrak{q}_2{}^{\Gamma\mp},\mathfrak{p}_2{}^{\Gamma\pm}]\subset\mathfrak{p}_{\Delta-\Gamma,1}{}^{\pm}$ by addition of eigenvalues of $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$ and because $[\mathfrak{k}^C,\mathfrak{p}^{\pm}]\subset\mathfrak{p}^{\pm}$. $[\mathfrak{p}_2{}^{\Gamma\pm},\mathfrak{p}_2{}^{\Gamma\pm}]=0$ and $[\mathfrak{q}_2{}^{\Gamma\mp},\mathfrak{q}_2{}^{\Gamma\mp}]=0$ now by Lemma 6.2. Finally $[\mathfrak{r}^{\Gamma\pm},\mathfrak{p}_{\Delta-\Gamma,1}{}^{\pm}]=0$ by addition of eigenvalues of $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$. Thus $\mathfrak{r}^{\Gamma\pm}$ is nilpotent of degree 2.

$\mathfrak{n}^{\Gamma\pm}$ is a real form of $\mathfrak{r}^{\Gamma\pm}$, and $\mathrm{ad}(\mathfrak{r}^{\Gamma\pm})$ is unipotent on $\mathfrak{g}^C$ by addition of eigenvalues, by Lemma 6.3. *Q.E.D.*

6.5. THEOREM. $\mathfrak{b}^\Gamma=\mathfrak{p}_\Gamma{}^0+\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\cdot(\mathfrak{k}^{\Gamma *}+\mathfrak{n}^{\Gamma-})$; $\mathfrak{b}^\Gamma$ *is the sum of the nonpositive eigenspaces of* $\mathrm{ad}(-Y_{\Delta-\Gamma}{}^0)$ *on* $\mathfrak{g}^0$ *and is the normalizer of* $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{n}^{\Gamma-}$ *in* $\mathfrak{g}^0$. $\mathfrak{e}^\Gamma$ *is the subalgebra* $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\cdot(\mathfrak{k}^\Gamma+\mathfrak{n}^{\Gamma-})$ *of* $\mathfrak{b}^\Gamma$.

Proof. The isotropy subalgebra of $\mathfrak{g}^C$ at x^Γ is $\mathrm{ad}(c_{\Delta-\Gamma})\cdot(\mathfrak{k}^C+\mathfrak{p}^+)$, which we decompose as

$$\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{k}^C+\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{q}_2{}^{\Gamma C} \\ +\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{p}_2{}^{\Gamma+}+\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{p}_{\Delta-\Gamma,1}{}^{+}+\mathfrak{p}_\Gamma{}^{+}.$$

As $\tau_{\Delta-\Gamma}\mathfrak{k}^\Gamma=\mathfrak{k}^\Gamma$ we have $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{k}^{\Gamma C}=\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{k}^{\Gamma C}$. Lemma 6.2 gives us

$$\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{r}_2{}^{\Gamma-}=\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{r}_2{}^{\Gamma+}\subset\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{q}_2{}^{\Gamma C}+\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{p}_2{}^{\Gamma+}.$$

Finally $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\mathfrak{p}_{\Delta-\Gamma,1}{}^{-}=\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{p}_{\Delta-\Gamma,1}{}^{+}$. Thus the isotropy subalgebra of $\mathfrak{g}^C$ at x^Γ contains $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\cdot(\mathfrak{k}^{\Gamma C}+\mathfrak{r}^{\Gamma-})$. Now Lemma 6.3 say that $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\cdot(\mathfrak{k}^{\Gamma *}+\mathfrak{n}^{\Gamma-})$ lies in the isotropy subalgebra of $\mathfrak{g}^0$ at x^Γ.

The isotropy subalgebra of $\mathfrak{g}^0$ at x^Γ has dimension $\dim. G^0-\dim. U^\Gamma$, and this is equal to $\dim.\mathfrak{k}+\frac{1}{2}\dim.\mathfrak{p}_{\Delta-\Gamma,1}$ by Theorem 5.7. Now

$$\dim.\mathfrak{k}+\tfrac{1}{2}\dim.\mathfrak{p}_{\Delta-\Gamma,1}=\dim.\mathfrak{k}^\Gamma+\dim.\mathfrak{q}_2{}^\Gamma+\dim.\mathfrak{n}_1{}^{\Gamma-} \\ =\dim.\mathfrak{k}^{\Gamma *}+\dim.\mathfrak{n}^{\Gamma-}=\dim.\mathrm{ad}(c_{\Delta-\Gamma})^{-1}(\mathfrak{k}^{\Gamma *}+\mathfrak{n}^{\Gamma-}).$$

The final assertion of the Theorem is proved. As $\mathfrak{p}_\Gamma{}^0\subset\mathfrak{b}^\Gamma$, as

$$\mathfrak{p}_\Gamma{}^0\cap\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\cdot\ (\mathfrak{k}^{\Gamma *}+\mathfrak{n}^{\Gamma-})=0,$$

and as $\dim.\mathfrak{p}_\Gamma{}^0=\dim. M_\Gamma=\dim. U^\Gamma-\dim. S^\Gamma$, it follows that

$$\mathfrak{b}^\Gamma=\mathfrak{p}_\Gamma{}^0+\mathrm{ad}(c_{\Delta-\Gamma})^{-1}\cdot\ (\mathfrak{k}^{\Gamma *}+\mathfrak{n}^{\Gamma-}),$$

which is our main assertion.

The eigenspace assertion now follows from Lemma 6.3, and the normalizer assertion is immediate. *Q.E.D.*

6.6. *Remarks on* $\mathfrak{b}^\Gamma$. The linear transformation $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$ is 0 on $\mathfrak{p}_\Gamma$ and $\pm i$ on $\mathfrak{p}_{\Delta-\Gamma,1}$; thus $[\mathfrak{p}_\Gamma,\mathfrak{p}_{\Delta-\Gamma,1}]$ is in the $(\pm i)$-eigenspace of $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$ on $\mathfrak{k}$, which is zero. We conclude that

$$[\mathfrak{g}_\Gamma{}^C,\mathfrak{g}_{\Delta-\Gamma,1}{}^C]=0. \tag{6.6.1}$$

Recall that

$$\mathfrak{p}_1{}^\Gamma=\mathfrak{p}_\Gamma+\mathfrak{p}_{\Delta-\Gamma,1},\quad \mathfrak{k}_1{}^\Gamma=[\mathfrak{p}_1{}^\Gamma,\mathfrak{p}_1{}^\Gamma],\ \mathfrak{k}_\Gamma=[\mathfrak{p}_\Gamma,\mathfrak{p}_\Gamma]$$

and $\mathfrak{k}_{\Delta-\Gamma,1}=[\mathfrak{p}_{\Delta-\Gamma,1}\mathfrak{p}_{\Delta-\Gamma,1}]$. With (6.6.1) this gives $\mathfrak{k}_1{}^\Gamma=\mathfrak{k}_\Gamma\oplus\mathfrak{k}_{\Delta-\Gamma,1}$ (direct sum of ideals); now it follows that

$$\mathfrak{k}_1{}^{\Gamma *}=\mathfrak{k}_\Gamma\oplus\mathfrak{k}_{\Delta-\Gamma,1}{}^{*}. \tag{6.6.2}$$

Now (6.6.2) and $\mathfrak{g}^\Gamma=\mathfrak{k}^\Gamma+\mathfrak{p}^\Gamma=\mathfrak{l}_2{}^\Gamma\oplus\mathfrak{g}_1{}^\Gamma$ yield

$$\mathfrak{k}^{\Gamma *}+\mathfrak{p}_\Gamma{}^0=\mathfrak{g}_\Gamma{}^0\oplus\mathfrak{l}_2{}^\Gamma\oplus\mathfrak{k}_{\Delta-\Gamma,1}{}^{*}. \tag{6.6.3}$$

The algebra (6.6.3) is a reductive subalgebra of $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$ which is a complement to $\mathfrak{n}^{\Gamma-}$ in $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{b}^\Gamma$. It follows that

(6.6.4) $\mathfrak{n}^{\Gamma-}$ is the nilradical of $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{b}^{\Gamma}$,

(6.6.5) $\mathfrak{g}_{\Gamma}^{0} \oplus \mathfrak{l}_2^{\Gamma} \oplus \mathfrak{k}_{\Delta-\Gamma,1}^{*}$ is a reductive complement to $\mathfrak{n}^{\Gamma-}$ in $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{b}^{\Gamma}$, and

(6.6.6) $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{b}^{\Gamma} = (\mathfrak{g}_{\Gamma}^{0} \oplus \mathfrak{l}_2^{\Gamma} \oplus \mathfrak{k}_{\Delta-\Gamma,1}^{*}) + \mathfrak{n}^{\Gamma-}$, semidirect sum.

$\mathfrak{e}^{\Gamma}$ denotes the isotropy subalgebra of $\mathfrak{g}^0$ at x^{Γ}. As above we see that

(6.6.7) $\mathfrak{n}^{\Gamma-}$ is the nilradical of $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{e}^{\Gamma}$,

(6.6.8) $\mathfrak{k}_{\Gamma} \oplus \mathfrak{l}_2^{\Gamma} \oplus \mathfrak{k}_{\Delta-\Gamma,1}^{*}$ is a reductive complement to $\mathfrak{n}^{\Gamma-}$ in $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{e}^{\Gamma}$, and

(6.6.9) $\mathrm{ad}(c_{\Delta-\Gamma})\mathfrak{e}^{\Gamma} = (\mathfrak{k}_{\Gamma} \oplus \mathfrak{l}_2^{\Gamma} \oplus \mathfrak{k}_{\Delta-\Gamma,1}^{*}) + \mathfrak{n}^{\Gamma-}$, semidirect sum.

6.7. In order to describe B^{Γ} we define

(6.7.1) $P_{\Delta-\Gamma,1}^{\pm} = \exp(\mathfrak{p}_{\Delta-\Gamma,1}^{\pm}) \subset G^C$ and $N_1^{\Gamma\pm} = \mathrm{ad}(c_{\Delta-\Gamma})G^0 \cap P_{\Delta-\Gamma,1}^{\pm}$,

(6.7.2) $R^{\Gamma\pm} = \exp(\mathfrak{r}^{\Gamma\pm}) \subset G^C$ and $N^{\Gamma\pm} = \mathrm{ad}(c_{\Delta-\Gamma})G^0 \cap R^{\Gamma\pm}$.

Lemma 6.4 says that every element of $\mathrm{ad}(\mathfrak{r}^{\pm})$ is a nilpotent linear transformation of $\mathfrak{g}^C$. Thus $P_{\Delta-\Gamma,1}^{\pm}$ and $R^{\Gamma\pm}$ are unipotent subgroups of G^C, and

$$\exp : \mathfrak{p}_{\Delta-\Gamma,1}^{\pm} \to P_{\Delta-\Gamma,1}^{\pm} \text{ and } \exp : \mathfrak{r}^{\Gamma\pm} \to R^{\Gamma\pm}$$

are one-one onto. In particular, the groups $P_{\Delta-\Gamma,1}^{\pm}$ and $R^{\Gamma\pm}$ are connected simply connected nilpotent Lie groups. Now let η be conjugation of G^C over $\mathrm{ad}(c_{\Delta-\Gamma})G^0$; η induces involutive automorphisms of the real groups $P_{\Delta-\Gamma,1}^{\pm}$ and $R^{\Gamma\pm}$ with respective fixed point sets $N_1^{\Gamma\pm}$ and $N^{\Gamma\pm}$. It follows that

(6.7.3) $N_1^{\Gamma\pm}$ and $N^{\Gamma\pm}$ are the analytic subgroups of $\mathrm{ad}(c_{\Delta-\Gamma})G^0$ with Lie algebras $\mathfrak{n}_1^{\Gamma\pm}$ and $\mathfrak{n}^{\Gamma\pm}$.

In particlar, $N_1^{\Gamma\pm}$ and $N^{\Gamma\pm}$ are connected simply connected nilpotent Lie groups.

6.8. THEOREM. *B^{Γ} is a parabolic subgroup of G^0 and is the normalizer of* $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}N^{\Gamma-}$ *in G^0. The identity component of B^{Γ} is given by*

$$B_0^{\Gamma} = \{G_{\Gamma}^{0} \cdot L_2^{\Gamma} \cdot \mathrm{ad}(c_{\Delta-\Gamma})^{-1}K_{\Delta-\Gamma,1}^{*}\} \cdot \mathrm{ad}(c_{\Delta-\Gamma})^{-1}N^{\Gamma-}$$

semidirect product; this is the Chevalley decomposition into reductive and unipotent parts.

Remark 1. *$G^0/B^{\Gamma} = S^{\Gamma}$ is a real projective variety defined over the rational number field.* For $B^{\Gamma} = B^{\Gamma C} \cap G^0$ for a parabolic subgroup $B^{\Gamma C}$ of G^C, $G^C/B^{\Gamma C}$ is a complex projective variety defined over the rationals, and a result of Borel [1, Proposition 3.7] gives the conjugation of G^C over G^0 defined over the rationals.

Remark 2. The reductive part of B_0^Γ admits

$$G_\Gamma^0 \times L_2^\Gamma \times \mathrm{ad}(c_{\Delta-\Gamma})^{-1}K_{\Delta-\Gamma,1}^*$$

as a covering group.

Remark 3. B^Γ is the subgroup of G^0 which preserves the boundary component $\xi^{-1}c_{\Delta-\Gamma}M_\Gamma = (\xi^{-1}c_{\Delta-\Gamma}\xi)D_\Gamma$ of D in $\mathfrak{p}^-$.

Proof of Theorem. Let $B^{\Gamma C}$ denote the analytic subgroup of G^C whose Lie algebra is the complexification $\mathfrak{b}^{\Gamma C}$ of $\mathfrak{b}^\Gamma$. As $-Y_{\Delta-\Gamma}^0$ is a basis of the Lie algebra of a split algebraic torus of G^0, and as $\mathfrak{b}^{\Gamma C}$ (resp $\mathfrak{b}^\Gamma$) is the sum of the nonpositive weight spaces of $\mathrm{ad}(-Y_{\Delta-\Gamma}^0)$ on $\mathfrak{g}^C$ (resp. on $\mathfrak{g}^0$), $B^{\Gamma C}$ is a parabolic subgroup of G^C and $B^{\Gamma C} \cap G^0$ is parabolic in G^0.

Let B_1^Γ denote $G^0 \cap B^{\Gamma C}$. Now B_0^Γ is the identity component both of B^Γ and B_1^Γ, and $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}N^{\Gamma-}$ is the unipotent radical of all three by (6.6.4) and (6.7.3). B_1^Γ is the full normaliser of $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}N^{\Gamma-}$ in G^0 because $B^{\Gamma C}$ is the full normaliser of $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}R^{\Gamma-}$ in G^C; now $B^\Gamma \subset B_1^\Gamma$. On the other hand $\mathfrak{b}^{\Gamma C} \subset \mathrm{ad}(c_{\Delta-\Gamma})(\mathfrak{k}^C + \mathfrak{p}^+)$ as in Theorem 6.5 so $B^{\Gamma C} \subset \mathrm{ad}(c_{\Delta-\Gamma})(K^C \cdot P^+)$; thus $B_1^\Gamma = B^{\Gamma C} \cap G^0 \subset \mathrm{ad}(c_{\Delta-\Gamma})((K^C \cdot P^+) \cap G^0) = B^\Gamma$. Now $B^\Gamma = B_1^\Gamma$, parabolic subgroup of G^0 which is the normalizer of $\mathrm{ad}(c_{\Delta-\Gamma})^{-1}N^{\Gamma-}$ in G^0.

The assertions on B_0^Γ now follow from Theorem 6.5, (6.6.5) and (6.6.6). *Q.E.D.*

6.9. COROLLARY. *If M is irreducible and of rank r, and if Γ has precisely t elements, then B^Γ is conjugate in G^0 to the group $F_{\Phi,}^0$ of Theorem* 3.4.

Remark 1. This corollary identifies B^Γ for reducible M by means of Corollary 4.10.

Remark 2. It is instructive to compare Corollary 6.9 with Theorem 4.13.

Proof of Corollary. Recall the maximally split Cartan subalgebra $\mathfrak{t} = \mathfrak{h}^- + \mathfrak{a}^0$ of $\mathfrak{g}^0$. Now $\mathfrak{u} = \mathfrak{h}^- + J\mathfrak{a}^0 = \mathrm{ad}(\exp(\pi/4)Z)\mathfrak{t}$ is a maximally split Cartan subalgebra of $\mathfrak{g}^0$; $J\mathfrak{a}^0$, the span of $\{Y_\delta^0\}_{\delta\in\Delta}$, is the split part of $\mathfrak{u}$.

$\Delta = \{\delta_1, \cdots, \delta_r\}$ with $\delta_1 < \cdots < \delta_r$. Define $\Delta(a)$ to be the last a elements of Δ so $\Gamma = \Delta(t)$. Let $Y(a) = -Y_{\Delta-\Delta(a)}^0$; $\{Y(1), \cdots, Y(r)\}$ is a basis of the split part $J\mathfrak{a}^0$ of $\mathfrak{u}$. We order the dual space of $J\mathfrak{a}^0$ lexicographically by values on this basis. Let β be a $\mathfrak{t}^C$-root of $\mathfrak{g}^C$. Then $\mathrm{ad}(c_\Delta)^*\beta = \beta^*$ is an $\mathfrak{h}^C$-root. If β^* is noncompact positive, then $\beta^*(Z_{\Delta-\Delta(a)}^0)$ is 0, $-i/2$ or $-i$; as

$$\mathrm{ad}(c_\Delta^{-1})(2iZ_{\Delta-\Delta(a)}^0) = \mathrm{ad}(c_{\Delta-\Delta(a)})^{-1}(2iZ_{\Delta-\Delta(a)}^0) = Y(a).$$

we then have $\beta(Y(a))$ equal to 0, 1 or 2. If β^* is a compact simple root we similarly have $\beta(Y(a))$ equal to 0 or -1, so $\beta(Y(a)) \geqq 0$ if β^* is compact negative. Now $\beta\,|_{J\mathfrak{a}^0} > 0$ implies either β^* is noncompact positive or β^* is compact negative. $\mathrm{ad}(c_{\Delta-\Delta(a)})\mathfrak{b}^{\Delta(a)C}$ contains every noncompact negative and every compact positive $\mathfrak{h}^C$-root space; thus the $\mathfrak{b}^{\Delta(a)C}$ are parabolic for the split torus $J\mathfrak{a}^0$. Now our assertion is the content of Theorem 3.4. *Q.E.D.*

7. The partial Cayley transforms of *D*. In this section we shall apply the partial Cayley transformation $c_{\Delta-\Gamma}$, where Γ is a subset of Δ of the type considered in Section 6, to the domain D embedded in $\mathfrak{p}^-$. It will turn out that the result of this transformation is a Siegel domain of type III, which we shall describe explicitly by determining the action of $\mathrm{ad}(c_{\Delta-\Gamma})B^\Gamma$ on $\mathfrak{p}^-$.

7.1. ν and ν^0 denote the conjugation of $\mathfrak{g}^C$ with respect to $\mathfrak{g}$ and $\mathfrak{g}^0$, respectively; $\langle\ ,\ \rangle$ denoting the Killing form, we define a positive definite Hermitian form by $\langle U,V\rangle_\nu = -\langle U,\nu V\rangle$ on $\mathfrak{g}^C$. The adjoint of a linear transformation $\mathrm{ad}(V)$ $(V\in\mathfrak{g}^C)$ with respect to this form is given by $\mathrm{ad}(V)^* = -\mathrm{ad}(\nu V)$ (cf. [7], § 6.1). We have $\mathfrak{p}^\pm = \mathfrak{p}_{\Delta-\Gamma,1}{}^\pm + \mathfrak{p}_2{}^{\Gamma\pm} + \mathfrak{p}_\Gamma{}^\pm$. ν is a complex antilinear map of $\mathfrak{p}^\pm$ onto $\mathfrak{p}^\mp$ preserving this direct decomposition.

For any $E\in\mathfrak{p}^-$ we denote by E_1, E_2 and E_3 the projections of E onto $\mathfrak{p}_{\Delta-\Gamma,1}{}^-$, $\mathfrak{p}_2{}^{\Gamma-}$ and $\mathfrak{p}_\Gamma{}^-$, respectively. So $E = E_1 + E_2 + E_3$.

By Lemma 6.3, $\mathfrak{n}_1{}^{\Gamma-}$ is a real form of $\mathfrak{p}_{\Delta-\Gamma,1}{}^-$. The terms "real," "imaginary," "Hermitian" will always refer to this real form. As in Section 4, we have $\mathfrak{o}^\Gamma = \xi^{-1}(c_{\Delta-\Gamma}(x)) = i\sum_{\alpha\in\Delta-\Gamma} E_{-\alpha}\in i\mathfrak{n}_1{}^{\Gamma-}$. By [7, Proposition 6.2] applied to the pair $(\mathfrak{g}_{\Delta-\Gamma}{}^0, \mathfrak{k}_{\Delta-\Gamma})$, the orbit $K_{\Delta-\Gamma,1}{}^*(-i\mathfrak{o}^\Gamma)$ is a self-dual cone in $\mathfrak{n}_1{}^{\Gamma-}$; we shall denote it by $\mathfrak{c}^\Gamma$.

7.2. LEMMA. *For all $U\in\mathfrak{p}_2{}^{\Gamma+}$, we have $\tau_{\Delta-\Gamma}(U) = -[U,\mathfrak{o}^\Gamma]$.*

Proof. First we show that, restricted to $\mathfrak{p}_2{}^{\Gamma C} + \mathfrak{q}_2{}^{\Gamma C}$, $\tau_{\Delta-\Gamma}$ and $\mathrm{ad}(X_{\Delta-\Gamma})$ coincide. By Lemma 6.3, $\mathfrak{p}_2{}^{\Gamma C} + \mathfrak{q}_2{}^{\Gamma C}$ is the sum of the (± 1)-eigenspaces of $\mathrm{ad}(Y_{\Delta-\Gamma}{}^0)$ on $\mathfrak{g}^C$. We have $X_{\Delta-\Gamma} = iX_{\Delta-\Gamma}{}^0 = -i\,\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)(Y_{\Delta-\Gamma}{}^0)$, and $\mathfrak{p}_2{}^{\Gamma C} + \mathfrak{q}_2{}^{\Gamma C}$ is invariant under $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$. It follows that $\mathfrak{p}_2{}^{\Gamma C} + \mathfrak{q}_2{}^{\Gamma C}$ is the sum of the $(\pm i)$-eigenspaces of $\mathrm{ad}(X_{\Delta-\Gamma})$. Now, if $\mathrm{ad}(X_{\Delta-\Gamma})U = \pm iU$, then $\tau_{\Delta-\Gamma}(U) = (\exp(\pi/2)\mathrm{ad}(X_{\Delta-\Gamma}))(U) = e^{\pm i\pi/2}U = \pm iU$, proving the assertion.

To prove the Lemma, let $U\in\mathfrak{p}_2{}^{\Gamma+}$. Then

$$\tau_{\Delta-\Gamma}(U) = -[U,X_{\Delta-\Gamma}] = -[U,i\sum_{\alpha\in\Delta-\Gamma}E_\alpha] - [U,i\sum_{\alpha\in\Delta-\Gamma}E_{-\alpha}] = -[U,\mathfrak{o}^\Gamma],$$

since $[U,E_\alpha] = 0$ for all $\alpha\in\Delta-\Gamma$, $\mathfrak{p}^+$ being abelian. *Q.E.D.*

Definitions. For all $W \in D_\Gamma$ we define the linear transformation $\mu(W) : \mathfrak{p}_2^{\Gamma-} \to \mathfrak{p}_2^{\Gamma-}$ by

$$\mu(W)U = \mathrm{ad}(W)\tau_{\Delta-\Gamma}\,\nu(U),$$

For all $V \in \mathfrak{p}_2^{\Gamma-}$ we define the linear function $f_V : \mathfrak{p}_\Gamma^- \to \mathfrak{p}_2^{\Gamma-}$ by

$$f_V(W) = (I + \mu(W))V.$$

Finally, for all $W \in D_\Gamma$ we define the vector-valued bilinear form $\Lambda_W : \mathfrak{p}_2^{\Gamma-} \times \mathfrak{p}_{2\Gamma}^- \to \mathfrak{p}_{\Delta-\Gamma,1}^-$ by

$$\Lambda_W(U, V) = -(i/2)[U, \tau_{\Delta-\Gamma}(\nu(I + \mu(W)))^{-1}V].$$

It is easy to see that these definitions are meaningful; for the definition of Λ_W we only have to note that $\|\mu(W)\| < 1$ in the operator norm with respect to the real part of $\langle\ ,\ \rangle_\nu$ restricted to $\mathfrak{p}_2^{\Gamma-}$. In fact, $\tau_{\Delta-\Gamma}$ and ν are isometric transformations on $\mathfrak{g}^C$ in this norm; $\tau_{\Delta-\Gamma}\,\nu$ maps $\mathfrak{p}_2^{\Gamma-}$ onto $\mathfrak{q}_2^{\Gamma+}$, and on $\mathfrak{q}_2^{\Gamma+}$ we have $\|\mathrm{ad}(W)\| < 1$ for all $W \in D_\Gamma$ by Lemma 4.6.

7.3. LEMMA.

(i) *For all $k \in K^{\Gamma*}$, $W \in D_\Gamma$ and $U, V \in \mathfrak{p}_2^{\Gamma-}$,*
$$\mathrm{ad}(k)\Lambda_W(U, V) = \Lambda_{\mathrm{ad}(k)W}(\mathrm{ad}(k)U, \mathrm{ad}(k)V).$$

(ii) *For all $W \in D_\Gamma$ and $U, V \in \mathfrak{p}_2^{\Gamma-}$,*
$$\Lambda_0(U, \mu(W)V)) = \Lambda_0(V, \mu(W)U).$$

(iii) *For all $W \in D_\Gamma$ we have $\Lambda_W = \Lambda_W^{(1)} + \Lambda_W^{(2)}$ where*

$$\Lambda_W^{(1)}, \Lambda_W^{(2)} : \mathfrak{p}_2^{\Gamma-} \times \mathfrak{p}_2^{\Gamma-} \to \mathfrak{p}_{\Delta-\Gamma,1}^-$$

are defined by

$$\Lambda_W^{(1)}(U, V) = -(i/2)[U, \tau_{\Delta-\Gamma}\,\nu\,(1 - \mu(W)^2)^{-1}V];$$
$$\Lambda_W^{(2)}(U, V) = (i/2)[U, \tau_{\Delta-\Gamma}\,\nu(1 - \mu(W)^2)^{-1}\mu(W)V].$$

$\Lambda_W^{(1)}$ is Hermitian bilinear and such that $\Lambda_W^{(1)}(U, U) \in \overline{\mathfrak{c}^\Gamma}$ for all $U \in \mathfrak{p}_2^{\Gamma-}$; $\Lambda_W^{(2)}$ is complex bilinear symmetric.

(iv) *For any $W \in D_\Gamma$, Λ_W is nondegenerate in the sense that if $\Lambda_W(U, V_0) = 0$ for all $U \in \mathfrak{p}_2^{\Gamma-}$, then $V_0 = 0$.*

(v) *For any fixed $U, V \in \mathfrak{p}_2^{\Gamma-}$, $\Lambda_W(U, f_V(W))$ is a constant vector, independent of W.*

Proof. Since $\|\mu(W)\| < 1$ for all $W \in D_\Gamma$, we have the convergent series expansions $(I + \mu(W))^{-1} = \sum_{n=0}^{\infty}(-\mu(W))^n$ and $(I - \mu(W)^2)^{-1} = \sum_{n=0}^{\infty}\mu(W)^{2n}$. These will be used several times in the proof.

To prove (i) we note that $\mathfrak{k}^{\Gamma*} = \mathfrak{l}^1 + i\mathfrak{q}_1^\Gamma$. $\tau_{\Delta-\Gamma}$ and ν are both trivial

on $\mathfrak{l}^\Gamma$ and equal to $-I$ on $i\mathfrak{q}_1{}^\Gamma$. Hence $\tau_{\Delta-\Gamma}\,\nu$ is trivial on $\mathfrak{k}^{\Gamma*}$, and thus $\mathrm{ad}(k)$ commutes with $\tau_{\Delta-\Gamma}\,\nu$ for all $k\in K^{\Gamma*}$. Also, $\mathrm{ad}(k)$ preserves $\mathfrak{p}_{\Delta-\Gamma,1}{}^-$, $\mathfrak{p}_2{}^{\Gamma-}$ and $\mathfrak{p}_{\Gamma}{}^-$ by Lemma 6.3.

It follows that

$$\begin{aligned}\mathrm{ad}(k)\Lambda_W(U,V) &= -(i/2)ad(k)\,[U,\tau_{\Delta-\Gamma}\,\nu\,(I+\mu(W))^{-1}V]\\ &= -(i/2)\,[\mathrm{ad}(k)U,\tau_{\Delta-\Gamma}\,\nu\,(I+\mu(\mathrm{ad}(k)W))^{-1}\,\mathrm{ad}(k)V]\\ &= \Lambda_{\mathrm{ad}(k)W}(\mathrm{ad}(k)U,\mathrm{ad}(k)V).\end{aligned}$$

To prove (ii) we use the definition of Λ_0, the fact that $\tau_{\Delta-\Gamma}$ commutes with ν (by definition of $\tau_{\Delta-\Gamma}$), then the Jacobi identity and the fact that $[U,V]=0$:

$$\begin{aligned}\Lambda_0(U,\mu(W)V) &= -(i/2)\,[U,\tau_{\Delta-\Gamma}\,\nu(V)]]\\ &= -(i/2)\,[U,[\tau_{\Delta-\Gamma}\,\nu(W),\ -V]]\\ &= -(i/2)\,[V,[\tau_{\Delta-\Gamma}\,\nu(W),\ -U]]\\ &= \Lambda_0(V,\mu(W)U).\end{aligned}$$

To prove (iii) we note that

$$\begin{aligned}\Lambda_W{}^{(1)}(U,V) &= \Lambda_0(U,\,(I-\mu(W)^2)^{-1}V),\\ \Lambda_W{}^{(2)}(U,V) &= \Lambda_0(U,\,(I-\mu(W)^2)^{-1}\mu(W)V).\end{aligned}$$

Hence $\Lambda_W=\Lambda_W{}^{(1)}+\Lambda_W{}^{(2)}$ is immediate.

Now we prove that $\Lambda_0=\Lambda_0{}^{(1)}$ is Hermitian bilinear and $\Lambda_0(U,U)\in\overline{\mathfrak{c}^\Gamma}$ for all $U\in\mathfrak{p}_2{}^{\Gamma-}$. Since Λ_0 is linear in the first and antilinear in the second argument, it suffices to show that $\Lambda_0(U,U)\in\overline{\mathfrak{c}^\Gamma}$ for all U. Since $\overline{\mathfrak{c}^\Gamma}$ is a self-dual cone, for this we only have to show that $\langle\Lambda_0(U,U),V\rangle_\nu\geqq 0$ for all $U\in\mathfrak{p}_2{}^{\Gamma-}$; $V\in\mathfrak{c}^\Gamma$.

Given any such U and V, there exists an element $k\in K_{\Delta-\Gamma,1}{}^*$ such that $\mathrm{ad}(k)V=-i\mathfrak{o}^\Gamma$. Denoting $U'=\mathrm{ad}(k)U$ and using (i) we have

$$\langle\Lambda_0(U,U),V\rangle_\nu=\langle\Lambda_0(U',U'),-i\mathfrak{o}^\Gamma\rangle_\nu.$$

Now note that by Lemma 7.2 we have

$$\tau_{\Delta-\Gamma}\,\nu(U')=-[\nu(U'),\mathfrak{o}^\Gamma]=-\mathrm{ad}(\nu(U'))\,\mathfrak{o}^\Gamma=\mathrm{ad}(U')^*\mathfrak{o}^\Gamma.$$

Hence

$$\Lambda_0(U',U')=-(i/2)\,[U',\tau_{\Delta-\Gamma}\,\nu(U')]=-(i/2)\cdot\mathrm{ad}(U')\mathrm{ad}(U')^*\mathfrak{o}^\Gamma.$$

Therefore,

$$\begin{aligned}\langle\Lambda_0(U,V),V\rangle_\nu &= \langle-(i/2)\ \mathrm{ad}(U')\mathrm{ad}(U')^*\mathfrak{o}^\Gamma,-i\mathfrak{o}^\Gamma\rangle_\nu\\ &= \tfrac{1}{2}\langle\mathrm{ad}(U')^*\mathfrak{o}^\Gamma,\mathrm{ad}(U')^*\mathfrak{o}^\Gamma\rangle_\nu\geqq 0,\end{aligned}$$

proving the assertion.

To prove the desired properties of $\Lambda_W^{(1)}$ for arbitrary $W \in D_\Gamma$, we first note that $\Lambda_0(U, \mu(W)^2 V) = \Lambda_0(\mu(W)V, \mu(W)U) = \Lambda_0(\mu(W)^2 U, V)$ for all U, V by (ii) and by hermiticity of Λ_0. Repeated application of these identities gives $\Lambda_0(U, \mu(W)^{2n}U) = \Lambda_0(\mu(W)^n U, \mu(W)^n U)$ for all $n \geqq 0$. Now we have

$$\Lambda_W^{(1)}(U,U) = \sum_{n=0}^{\infty} \Lambda_0(U, \mu(W)^{2n}U) = \sum_{n=0}^{\infty} \Lambda_0(\mu(W)^n U, \mu(W)^n U).$$

By what we just proved, each term of the last sum is in $\overline{\mathfrak{c}^\Gamma}$; hence $\Lambda_W(U,U) \in \overline{\mathfrak{c}^\Gamma}$ for all $U \in \mathfrak{p}_2^{\Gamma-}$. Since $\Lambda_W^{(1)}$ is linear in the first, antilinear in the second argument (by complex linearity of $\mu(W)^2$), this also shows that $\Lambda_W^{(1)}$ is Hermitian, as we had to prove.

$\Lambda_W^{(2)}$ is clearly complex bilinear for any $W \in D_\Gamma$. To prove that it is symmetric, we use the definition of $\Lambda_W^{(2)}$, hermiticity of $\Lambda_0(U, (I - \mu(W)^2)^{-1}V) = \Lambda_W^{(1)}(U, V)$ in U and V, hermiticity of Λ_0, then (ii) and again the definition of $\Lambda_W^{(2)}$; denoting the conjugation of $\mathfrak{p}_{\Delta-\Gamma,1}^-$ with respect to $\mathfrak{n}_1^{\Gamma-}$ by ρ, we have

$$\begin{aligned}\Lambda_W^{(2)}(U,V) &= \Lambda_0(U, (I - \mu(W)^2)^{-1}\mu(W)V) \\ &= \rho\Lambda_0(\mu(W)V, (I - \mu(W)^2)^{-1}U) \\ &= \Lambda_0((I - \mu(W)^2)^{-1}U, \mu(W)V) \\ &= \Lambda_0(V, \mu(W)(I - \mu(W)^2)^{-1}U) \\ &= \Lambda_W^{(2)}(V, U).\end{aligned}$$

This finishes the proof of (iii).

In order to prove (iv) it is enough to show that Λ_0 is non-degenerate. The relation $\Lambda_W(U,V) = \Lambda_0(U, (I + \mu(W))^{-1}V)$ will then imply that Λ_W is non-degenerate. We show that $\Lambda_0(U,U) = 0$ implies $U = 0$.

Suppose $\Lambda_0(U,U) = 0$ for some $U \in \mathfrak{p}_2^{\Gamma-}$. As in the proof of (iii), we have by Lemma 7.2,

$$0 = \langle \Lambda_0(U,U), -i\mathfrak{o}^\Gamma \rangle_\nu = \tfrac{1}{2}\langle \mathrm{ad}(U)^*\mathfrak{o}^\Gamma, \mathrm{ad}(U)^*\mathfrak{o}^\Gamma \rangle_\nu.$$

Since $\langle\ ,\ \rangle_\nu$ is positive definite, this implies $\mathrm{ad}(U)^*\mathfrak{o}^\Gamma = 0$. This means $[-\nu(U), \mathfrak{o}^\Gamma] = 0$. Now, since $\mathfrak{o}^\Gamma = \sum_{\alpha \in \Delta-\Gamma} E_{-\alpha}$, it follows that $[\nu(U), Y_{\Delta-\Gamma}{}^0] = 0$, i.e. $\mathrm{ad}(Y_{\Delta-\Gamma}{}^0(\nu(U)) = 0$. Since $\nu(U) \in \mathfrak{p}_2^{\Gamma+}$, by Lemma 6.3 it follows that $\nu(U) = 0$. Hence $U = 0$, as we had to show.

The proof of (v) is trivial from the definitions; we have

$$\begin{aligned}\Lambda_W(U, f_V(W)) &= -(i/2)[U, \tau_{\Delta-\Gamma}\,\nu(I + \mu(W))^{-1}(I + \mu(W))V] \\ &= -(i/2)[U, \tau_{\Delta-\Gamma}\,\nu(V)],\end{aligned}$$

which is independent of W. *Q.E.D.*

7.4. LEMMA. *The map $I - \tau_{\Delta-\Gamma}\,\nu$ is a real linear isomorphism of $\mathfrak{p}_2^{\Gamma-}$*

onto $\mathfrak{n}_2^{\Gamma-}$; *so every element of* $\mathfrak{n}_2^{\Gamma-}$ *can uniquely be written as* $V-\tau_{\Delta-\Gamma}\,\nu(V)$, *with* $V\in\mathfrak{p}_2^{\Gamma-}$.

Proof. If $V\in\mathfrak{p}_2^{\Gamma-}$, then $\tau_{\Delta-\Gamma}\,\nu(V)\in\mathfrak{q}_2^{\Gamma+}$. $\mathfrak{n}_2^{\Gamma-}=(\mathfrak{p}_2^{\Gamma-}+\mathfrak{q}_2^{\Gamma+})\cap\operatorname{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$ is a real form of $\mathfrak{p}_2^{\Gamma-}+\mathfrak{q}_2^{\Gamma+}$, hence dim. $\mathfrak{n}_2^{\Gamma-}=$ dim. $\mathfrak{p}_2^{\Gamma-}=$ dim. $\mathfrak{q}_2^{\Gamma+}$. Now it suffices to prove that $V-\tau_{\Delta-\Gamma}\,\nu(V)\in\operatorname{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$ for all $V\in\mathfrak{p}_2^{\Gamma-}$.

The involution of $\mathfrak{g}^C$ with respect to $\operatorname{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$ is

$$\begin{aligned}\operatorname{ad}(c_{\Delta-\Gamma})\nu^0\operatorname{ad}(c_{\Delta-\Gamma})^{-1}&=\operatorname{ad}(c_{\Delta-\Gamma})\sigma\nu\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\\&=\sigma\operatorname{ad}(c_{\Delta-\Gamma})^{-1}\nu\operatorname{ad}(c_{\Delta-\Gamma})^{-1}=\sigma\nu\tau_{\Delta-\Gamma}^{-1}.\end{aligned}$$

For $V\in\mathfrak{p}_2^{\Gamma-}$ we have

$$\begin{aligned}\sigma\nu\tau_{\Delta-\Gamma}^{-1}(V)&=-\sigma\nu\tau_{\Delta-\Gamma}(V)=-\tau_{\Delta-\Gamma}\,\nu(V),\\ \sigma\nu\tau_{\Delta-\Gamma}^{-1}(\tau_{\Delta-\Gamma}\,\nu(V))&=\sigma(V)=-V.\end{aligned}$$

Hence $V-\tau_{\Delta-\Gamma}\,\nu(V)$ is invariant under $\sigma\nu\tau_{\Delta-\Gamma}^{-1}$, and so is contained in $\operatorname{ad}(c_{\Delta-\Gamma})\mathfrak{g}^0$. *Q.E.D.*

7.5. PROPOSITION. $N^{\Gamma-}$ *acts on* $\mathfrak{p}^-$ *by*

$$g(E)=E+U+f_V(E_3)+2i\Lambda_{E_2}(E_2,f_V(E_3))+i\Lambda_{E_3}(f_V(E_3),f_V(E_3))$$

where $g=\exp(U+(I-\tau_{\Delta-\Gamma}\,\nu)(V))$, $U\in\mathfrak{n}_1^{\Gamma-}$, $V\in\mathfrak{p}_2^{\Gamma-}$. $K^{\Gamma *}$ *acts on* $\mathfrak{p}^-$ *by the adjoint representation; it preserves* $\mathfrak{p}_{\Delta-\Gamma,1}^-$, $\mathfrak{p}_2^{\Gamma-}$ *and* $\mathfrak{p}_\Gamma^-$. *On* $\mathfrak{p}_{\Delta-\Gamma,1}^-$, $K_{\Delta-\Gamma,1}{}^*$ *is real,* K_Γ *and* L_2^Γ *are trivial. These actions are* ξ*-equivariant; in particular* $K^{\Gamma *}\cdot N^-$ *preserves* $\xi(\mathfrak{p}^-)$.

Proof. It is easy to see that K^C and P^- act on $\mathfrak{p}^-$ in a ξ-equivariant way by the adjoint representation and by translations, respectively. Now let g be any element of $N^{\Gamma-}$; it can be written in the given form by Lemma 7.4. By the Campbell-Hausdorff formula we have

$$\begin{aligned}g&=\exp(U+V-\tau_{\Delta-\Gamma}\,\nu(V))\\&=\exp(U)\cdot\exp(\tfrac{1}{2}[V,\tau_{\Delta-\Gamma}\,\nu(V)])\cdot\exp(V)\cdot\exp(-\tau_{\Delta-\Gamma}\,\nu(V))\end{aligned}$$

since, by Lemmas 6.3 and 6.4, all other brackets vanish. Now $-\tau_{\Delta-\Gamma}\,\nu(V)\in\mathfrak{q}_2^{\Gamma+}\subset\mathfrak{k}^C$, so $\exp(-\tau_{\Delta-\Gamma}\,\nu(V))$ acts on $\mathfrak{p}^-$ by the adjoint action;

$$\begin{aligned}&\exp(-\tau_{\Delta-\Gamma}\,\nu(V))(E)\\&\quad=E-[\tau_{\Delta-\Gamma}\,\nu(V),E]+\tfrac{1}{2}[\tau_{\Delta-\Gamma}\,\nu(V),[\tau_{\Delta-\Gamma}\,\nu(V),E]]\\&\quad=E+[E_2,\tau_{\Delta-\Gamma}\,\nu(V)]+[E_3,\tau_{\Delta-\Gamma}\,\nu(V)]\\&\qquad+\tfrac{1}{2}[[E_3,\tau_{\Delta-\Gamma}\,\nu(V)],\tau_{\Delta-\Gamma}\,\nu(V)],\end{aligned}$$

since all other brackets vanish, again by Lemmas 6.3 and 6.4. The other factors in the expression of g are in P^-, so they act on $\mathfrak{p}^-$ by translations. Using the definition of Λ_W the assertion about the action of g follows.

$K^{\Gamma *}$ commutes with $\mathrm{ad}(Z_{\Delta-\Gamma}{}^0)$, therefore preserves its eigenspaces $\mathfrak{p}_{\Delta-\Gamma,1}{}^-$, $\mathfrak{p}_2{}^{\Gamma -}$ and $\mathfrak{p}_\Gamma{}^-$ in $\mathfrak{p}^-$. $K_{\Delta-\Gamma,1}{}^*$ is real on $\mathfrak{p}_{\Delta-\Gamma,1}{}^-$ by [7, Proposition 6.6] applied to the pair $(\mathfrak{g}_{\Delta-\Gamma}{}^0, \mathfrak{k}_{\Delta-\Gamma})$. $\mathfrak{l}_2{}^\Gamma$ and $\mathfrak{k}_\Gamma$ centralize $\mathfrak{p}_{\Delta-\Gamma,1}{}^-$, hence $L_2{}^\Gamma$ and K_Γ act trivially on it. *Q.E.D.*

7.6. We define the partial Cayley transform of D by

$$c_{\Delta-\Gamma}D = \xi^{-1}(c_{\Delta-\Gamma}G^0(x)).$$

This is the image of D under $\xi^{-1}c_{\Delta-\Gamma}\xi$ in $\mathfrak{p}^-$. To see that this definition is meaningful, we note that by (6.6.6) we have a local semidirect product $\mathrm{ad}(c_{\Delta-\Gamma})B_0{}^\Gamma = G_\Gamma{}^0\cdot(K_{\Delta-\Gamma,1}{}^*\cdot L_2{}^\Gamma N^{\Gamma -})$. Using this, we have

$$\begin{aligned} c_{\Delta-\Gamma}(G^0(x)) &= c_{\Delta-\Gamma}(B_0{}^\Gamma(x)) = (\mathrm{ad}(c_{\Delta-\Gamma})B_0{}^\Gamma)(c_{\Delta-\Gamma}x) \\ &= (N^{\Gamma -}\cdot L_2{}^\Gamma\cdot K_{\Delta-\Gamma,1}{}^*)(G_\Gamma{}^0(c_{\Delta-\Gamma}x)) \\ &= (N^{\Gamma -}\cdot L_2{}^\Gamma\cdot K_{\Delta-\Gamma,1}{}^*)(c_{\Delta-\Gamma}M_\Gamma) \subset \xi(\mathfrak{p}^-), \end{aligned}$$

by Lemma 4.2 and Proposition 7.5.

7.7. THEOREM. *The partial Cayley transform $c_{\Delta-\Gamma}D$ of D is the domain* $\{E : \mathrm{Im}\, E_1 - \mathrm{Re}\, \Lambda_{E_3}(E_2, E_2) \in \mathfrak{c}^\Gamma, E_3 \in D_\Gamma\}$.

Proof. Let us denote by S the domain defined in the text of the Theorem. First we show that $c_{\Delta-\Gamma}D \subset S$. By 7.6 we have

$$c_{\Delta-\Gamma}D = (N^{\Gamma -}\cdot L_2{}^\Gamma\cdot K_{\Delta-\Gamma,1}{}^*)(G_\Gamma{}^0(\mathfrak{o}^\Gamma)).$$

Now $G_\Gamma{}^0(\mathfrak{o}^\Gamma) = \{\mathfrak{o}^\Gamma + E_3 : E_3 \in D_\Gamma\} \subset S$, and in order to see that $c_{\Delta-\Gamma}D \subset S$, it suffices to show that $L_2{}^\Gamma\cdot K_{\Delta-\Gamma,1}{}^*$ and $N^{\Gamma -}$ map S into itself. To show this, let $E \in S$ and let $k \in L_2{}^\Gamma\cdot K_{\Delta-\Gamma,1}{}^*$. We denote $E' = \mathrm{ad}(k)E$. By Lemma 7.3(i) and Proposition 7.5 we have

$\mathrm{Im}\, E'_1 - \mathrm{Re}\, \Lambda_{E_3}(E'_2, E'_2) = \mathrm{ad}(k)\,(\mathrm{Im}\, E_1 - \mathrm{Re}\, \Lambda_{E_3}(E_2, E_2)) \in \mathrm{ad}(k)\mathfrak{c}^\Gamma = \mathfrak{c}^\Gamma$.
Now let $g = \exp(U + (I - \tau_{\Delta-\Gamma}\,\nu)(V)) \in N^{\Gamma -}$ with $U \in \mathfrak{n}_1{}^{\Gamma -}$, $V \in \mathfrak{p}_2{}^{\Gamma -}$, and let $E' = g(E)$. By Proposition 7.5 we have

$$\begin{aligned} &\mathrm{Im}\, E'_1 - \mathrm{Re}\, \Lambda_{E'_3}(E'_2, E'_2) \\ &\quad = \mathrm{Im}(E_1 + U + 2i\Lambda_{E_3}(E_2, f_V(E_3)) + i\Lambda_{E_3}(f_V(E_3), f_V(E_3))) \\ &\qquad\qquad - \mathrm{Re}\, \Lambda_{E_3}(E_2 + f_V(E_3), E_2 + f_V(E_3))) \\ &\quad = \mathrm{Im}\, E_1 - \mathrm{Re}\, \Lambda_{E_3}(E_2, E_2), \end{aligned}$$

proving the assertion.

Next we prove that $S \subset c_{\Delta-\Gamma}D$. Let $E \in S$, it is sufficient to show that E can be transformed into the element $\mathfrak{o}^\Gamma \in c_{\Delta-\Gamma}D$ by an element of $\mathrm{ad}(c_{\Delta-\Gamma})B_0{}^\Gamma$. Let $V = -(I + \mu(E_3))^{-1}E_2$; then

$$n_2 = \exp((I - \tau_{\Delta-\Gamma}\,\nu)(V)) \in N^{\Gamma -}.$$

carries E into an element $E' = E'_1 + 0 + E_3$. Now let $U = -\operatorname{Re} E'_1$; then $n_1 = \exp(U) \in N^{\Gamma-}$ carries E' into $E'' = iF + 0 + E_3$, with F real. As we showed above, $N^{\Gamma-}$ preserves S, so we have

$$\operatorname{Im} E''_1 - \operatorname{Re} \Lambda_{E''_3}(E''_2, E''_2) = F \in \mathfrak{c}^{\Gamma}.$$

Now there exists an element $k \in K_{\Delta-\Gamma,1}{}^{*}$ such that $k \cdot F = -i\mathfrak{o}^{\Gamma}$; k carries E'' into $E''' = \mathfrak{o}^{\Gamma} + 0 + E_3$. Finally, since $E_3 \in D_{\Gamma}$, there exists $g \in G_{\Gamma}{}^{0}$ such that $g \cdot E_3 = 0$. It follows that $gkn_1n_2 \cdot E = \mathfrak{o}^{\Gamma}$, and $gkn_1n_2 \in \operatorname{ad}(c_{\Delta-\Gamma})B^{\Gamma}$.
Q.E.D.

UNIVERSITY OF CALIFORNIA, BERKELEY.

REFERENCES.

[1] A. Borel, "Compact Clifford-Klein forms of symmetric spaces," *Topology*, vol. 2 (1963), pp. 111-122.

[2] A. Borel and J. Tits, "Groupes réductifs," *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, No. 27 (1965), pp. 55-150.

[3] H. Furstenberg, "A Poisson formula for semisimple Lie groups," *Annals of Mathematics* (2), vol. 77 (1963), pp. 335-386.

[4] Harish-Chandra, "Representations of semisimple Lie groups, VI," *American Journal of Mathematics*, vol. 78 (1956), pp. 564-628.

[5] R. Hermann, "Geometric aspects of potential theory in bounded symmetric domains I," *Mathematische Annalen*, vol. 148 (1962), pp. 349-366; II, *ibid.*, vol. 151 (1963), pp. 143-149; III, *ibid.*, vol. 153 (1964), pp. 384-394.

[6] L. K. Hua and K. H. Look, "Theory of harmonic functions in classical domains," *Scientia Sinica*, vol. 8 (1959), pp. 1031-1094.

[7] A. Korányi and J. A. Wolf, "Realization of hermitian symmetric spaces as generalized half-planes," *Annals of Mathematics* (2), vol. 81 (1965), pp. 265-288.

[8] C. C. Moore, "Compactifications of symmetric spaces II (The Cartan domains)," *American Journal of Mathematics*, vol. 86 (1964), pp. 358-378.

[9] I. I. Pjateckiĭ-Šapiro, *Geometry of classical domains and theory of automorphic functions*, Moscow, 1961 (in Russian).

[10] I. Satake, "On representations and compactifications of symmetric Riemannian spaces," *Annals of Mathematics* (2), vol. 71 (1960), pp. 77-110.

[11] J. Tits, "Espaces homogènes complexes compacts," *Commentarii Mathematici Helvetici*, vol. 37 (1962), pp. 111-120.

[12] J. A. Wolf, "Self adjoint function spaces on Riemannian symmetric manifolds," *Transactions of the American Mathematical Society*, vol. 113 (1964), pp. 299-315.

A MÜNTZ-JACKSON THEOREM.

By D. J. Newman.

Let $0 = \lambda_0 < \lambda_1 < \lambda_2 \cdots$ be integers. The beautiful theorem of Müntz [2] states that the monomials $\{x^\lambda\}$, $\lambda = \lambda_0, \lambda_1, \cdots$, span all of $C[0.1]$ if and only if $\sum 1/\lambda_i = \infty$. This generalizes the famous theorem of Weierstrass that the monomials $1, x, x^2, x^3, \cdots$, span all of $C[0,1]$.

There is a rather different kind of generalization of the Weierstrass result embodied in the theorem of Jackson [1], p. 36. This states that for any $f(x)$, there exist $C_0, C_1 \cdots C_n$ such that

$$| f(x) - \sum_{i=0}^{n} C_i x^i | \leqq 6\omega_f(1/n)$$

throughout $[0,1]$; $\omega_f(\delta)$ representing, as usual, the modulus of continuity of $f(x)$.

In this paper we attempt to combine these two types of generalization. In short we seek a version of Müntz' theorem which expresses the "speed" of convergence of the polynomials $\sum C_i x^\lambda$ to a function $f(x)$, where $\lambda = \lambda_i$.

Unfortunately, however, we are at present only able to give results in the L^2 norm. Thus although our previous remarks refer to continuous functions and uniform convergence we must reorient ourselves to L^2 functions and L^2 convergence. To do so, let us recall the notion of the L^2 modulus of continuity.

Definition. For $f(x) \in L^2[0,1]$, we continue $f(x)$ to have period 1 and we denote $\omega_f(\delta) = \sup_{|t| \leqq \delta} [\int_0^1 |f(x+t) - f(x)|^2 \, dx]^{\frac{1}{2}}$.

We denote by S the class of functions $f(x)$ for which $\omega_f(\delta) \leqq \delta$.

Now let $0 = \lambda_0 < \lambda_1 \cdots < \lambda_n$ be a finite set Λ of integers. *We will assume that* $\lambda_{i+1} - \lambda_i > 1$. By $C(\Lambda)$ we mean all $\phi(x)$ with $L^2[0,1]$ norm one which are orthogonal to all x^λ, $\lambda \in \Lambda$. Finally, put

$$\epsilon_\Lambda = \prod_{i=1}^{n} (\lambda_i - \tfrac{1}{2})/(\lambda_i + \tfrac{3}{2}).$$

Our main theorem is as follows

Received December 21, 1964.
Revised March 21, 1965.

THEOREM 1. *Let $f(x) \in L^2[0,1]$, then there exist C_λ such that*

$$\| f(x) - \sum_{\lambda \in \Lambda} C_\lambda x^\lambda \| \leqq 3\omega_f(\epsilon_\Lambda).$$

We will also show that the right hand estimate is essentially best possible. For now, however, let us content ourselves with noting that Theorem 1 does indeed contain the L^2 versions of the Jackson and Müntz theorems. Indeed if $\sum_{i=1}^{\infty} 1/\lambda_i = \infty$, then, with $\Lambda = \{\lambda_1, \cdots, \lambda_n\}$, ϵ_Λ goes to 0 as $n \to \infty$, this gives the Müntz theorem. On the other hand by restricting $\lambda_i = 2i$, $i = 0, 1, \cdots, n$, we find $\epsilon_\Lambda = 3/(4n+3)$ which gives us a slightly stronger form of Jackson's theorem.

It is amusing to examine the case of the primes, $\lambda_i =$ the i-th prime $= p(i)$. For this sequence, Theorem 1 guarantees the existence of C_i for which $\| f(x) - C - \sum_{i=1}^{n} C_i x^{p(i)} \| \leqq C\omega_f(1/\log^2 n)$. And as we have previously remarked, this is best possible.

Let us now introduce the "approximation index," p_Λ, defined by

$$p_\Lambda = \sup_{f(x) \in S} \inf \| f(x) - \sum_{\lambda \in \Lambda} C_\lambda x^\lambda \|,$$

where the infimum is taken over all sets of constants C_λ. We will prove

THEOREM 2. $\frac{1}{4}\epsilon_\Lambda \leqq p_\Lambda \leqq \epsilon_\Lambda$.

Theorem 2 clearly establishes Theorem 1 in the case of $f \in S$ and also points out the sense in which it is best possible. We will show in fact, by an elementary argument, that Theorem 2 implies Theorem 1 in general. First, however, we turn to the proof of Theorem 2.

LEMMA 1. *$f(x) \in S$ if and only if $f(x)$ is equivalent (i. e., equal almost everywhere) to an absolutely continuous function $F(x)$ of period 1 with $\| F'(x) \| \leqq 1$.*

Proof. First assume that $F(x)$ is periodic, absolutely continuous, and $\| F' \| \leqq 1$, then

$$\int_0^1 | F(x+t) - F(x) |^2 \, dx = \int_0^1 \left| \int_x^{x+t} F'(u) \, du \right|^2 dx$$

$$\leqq \int_0^1 \int_x^{x+t} dv \int_x^{x+t} | F'(u) |^2 \, du dx = t \int_0^1 \int_x^{x+t} | F'(u) |^2 \, du dx$$

$$= t \int_0^1 | F'(u) |^2 \int_{u-t}^{u} dx du \leqq t^2 \| F' \|^2 \leqq t^2.$$

It follows that $F \in S$.

Now suppose $F(x) \in S$. This means that the set of functions

$$f_t(x) = (f(x+t) - f(x))/t$$

has L_2 norm uniformly bounded by 1. We may extract a weakly convergent subsequence with $t_n \to \infty$ which converges to $\phi(x)$, say. Clearly $\| \phi(x) \| \leqq 1$, and also by weak convergence,

$$\int_a^b (f(x+t_n) - f(x))/t_n \, dx \to \int_a^b \phi(x)\, dx.$$

The left side, however, approaches $f(b) - f(a)$ for almost all a and b, and the lemma is thereby proved.

LEMMA 2. $$p_\Lambda = \sup_{\phi \in C(\Lambda)} \| \int_0^x \phi(t)\, dt \|.$$

Proof. $C(\Lambda)$ is the portion of the unit sphere which lies in the orthogonal complement of the span, M, of $\{x^\lambda\}$, $\lambda \in \Lambda$. Let $f(x) \in S$. The distance from f to M is therefore given by $\sup(f, \phi)$. Lemma 1 allows for an integration by parts and we obtain $(f, \phi) = -(f'(x), \int_0^x \phi(t)\, dt)$. Again, by Lemma 1, the supremum, with respect to f of the left hand side is exactly $\| \int_0^x \phi(t)\, dt \|$. The lemma now follows easily.

Next we make use of the Paley-Wiener theory of Fourier transforms; see [2], pp. 1-13. Viz. for $\phi \in C(\Lambda)$ we set $F(z) = \int_0^1 t^{-(iz+\frac{1}{2})} \phi(t)\, dt$ and note that $F(z)$ is in the Paley-Wiener class for the upper half plane (hereafter called the class P). We note that

(1) $$F(i(\lambda + \tfrac{1}{2})) = 0 \text{ for all } \lambda \in \Lambda$$

and by Parseval's theorem,

(2) $$\frac{1}{2\pi} \int_{-\infty}^{\infty} | F(x) |^2 \, dx = \int_0^1 | \phi(t) |^2 \, dt = 1.$$

Next, integration by parts gives

$$F(z) = (iz + \tfrac{1}{2}) \int_0^1 t^{-(iz+\frac{3}{2})} [\int_0^t \phi(u)\, du]\, dt,$$

$$F(x+i) = (ix - \tfrac{1}{2}) \int_0^1 t^{-(ix+\frac{1}{2})} [\int_0^t \phi(u)\, du]\, dt$$

So that, again by Parseval's theorem

(3) $$\frac{1}{2\pi} \int_{-\infty}^{\infty} | F(x+i) |^2 / (x^2 + \tfrac{1}{4})\, dx = \int_0^1 | \int_0^t \phi(u)\, du |^2 \, dt.$$

Now write $B(z)=\prod[z-i(\lambda+\frac{1}{2})]/[z+i(\lambda+\frac{1}{2})]$, $\lambda\in\Lambda$, and $G(z)=F(z)/B(z)$. By (1) and the fact that $|B(x)|=1$, it follows that $G(z)\in P$, and that (2) becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}|G(x)|^2\,dx=1, \tag{4}$$

while from (3) we obtain

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}|G(x+i)|^2H(x)\,dx=\int_0^1\left|\int_0^t\phi(u)\,du\right|^2dt \tag{5}$$

where

$$H(x)=(x^2+\tfrac{1}{4})^{-1}\prod_{\lambda\in\Lambda}[x^2+(\lambda-\tfrac{1}{2})^2]/[x^2+(\lambda+\tfrac{3}{2})^2].$$

From (4) and (5) and Lemma 2, we conclude

LEMMA 3. *We have*

$$p_\Lambda^2=\sup_{G\in P}\int_{-\infty}^{\infty}|G(x+i)|^2H(x)\,dx\Big/\int_{-\infty}^{\infty}|G(x)|^2\,dx.$$

We can now prove Theorem 2. For the upper estimate note that

$$H(x)=\epsilon_\Lambda^2(1+x^2/(\lambda_n+\tfrac{3}{2})^2)^{-1}\prod_{i=1}^{n}[1+x^2/(\lambda_i-\tfrac{1}{2})^2]/[1+x^2/(\lambda_{i-1}+\tfrac{3}{2})^2].$$

If we recall the fact, then, that $\lambda_i-\frac{1}{2}\geqq\lambda_{i-1}+\frac{3}{2}$. We may conclude that $H(x)\leqq\epsilon_\Lambda^2$ for all x. Hence

$$p_\Lambda^2\leqq\epsilon_\Lambda^2\sup_{g\in P}\int_{-\infty}^{\infty}|G(x+i)|^2\,dx\Big/\int_{-\infty}^{\infty}|G(x)|^2\,dx$$

and the required inequality follows from the fact that

$$\int_{-\infty}^{\infty}|G(x+i)|^2\,dx\leqq\int_{-\infty}^{\infty}|G(x)|^2\,dx \text{ for all } G\in P.$$

For the lower bound note that

$$\begin{aligned}H(x)&=(x^2+9/4)^{-1}\prod_{i=1}^{n}[x^2+(\lambda_i-1/2)^2]/[x^2+(\lambda_i+3/2)^2]\\&\geqq(x^2+9/4)^{-1}\epsilon_\Lambda^2.\end{aligned}$$

Choosing $G(z)=(z+\frac{1}{2})^{-1}$, we obtain

$$p_\Lambda^2\geqq\epsilon_\Lambda^2\int_{-\infty}^{\infty}(x^2+9/4)^{-2}\,dx\Big/\int_{-\infty}^{\infty}(x^2+1/4)^{-1}\,dx=2\epsilon_\Lambda^2/27.$$

The proof is complete.

We conclude with the proof that Theorem 2 implies Theorem 1. Let $f(x)$ be given with the L^2 modulus of continuity $\omega(\delta)$. If $f(x)$ has the Fourier series $\sum_{-\infty}^{\infty} C_n e^{2\pi i n x}$ then

$$\omega^2(\delta) = \sup_{|t|\leqq\delta} \int_0^1 \left| \sum C_n (e^{2\pi i n t} - 1) e^{2\pi i n x} \right|^2 dx \tag{6}$$
$$= \sup_{|t|\leqq\delta} 4 \sum |C_n|^2 \sin^2 \pi n t.$$

From (6) two lower estimates emerge. First

$$\omega^2(\delta) \geqq 4 \sum |C_n|^2 \sin^2 \pi n\delta \geqq 16\delta^2 \sum_{|n|\leqq 1/2\delta} |C_n|^2 n^2 \tag{7}$$
$$= 4\delta^2/\pi^2 \sum_{|n|\leqq 1/2\delta} |C_n|^2 (2\pi n)^2,$$

while also

$$\omega^2(\delta) \geqq (4/\delta) \int_0^{\delta} \sum |C_n|^2 \sin^2 \pi n t \, dt = 2 \sum |C_n|^2 [1 - \sin 2\pi n\delta / 2\pi n\delta] \tag{8}$$
$$\geqq \sum |C_n|^2 (2 - 1/\delta\pi n) \geqq \sum_{|n|>\frac{1}{2}\delta} |C_n|^2.$$

Let $F(x)$ denote the $(1/2\delta)$—partial sum of the Fourier series or $f(x)$. We conclude from (7) that $\| F'(x) \|^2 \leqq \pi^2\omega^2(\delta)/4\delta^2$ or, by Lemma 1, that

$$[2\delta/\pi\omega(\delta)] F(x) \in S. \tag{9}$$

and we can by Theorem 2, determine C_λ such that

$$\| F(x) - \sum_{x \in \Lambda} C_\lambda x^\lambda \| \leqq [\pi\omega(\delta)/2\delta] \epsilon_\Lambda. \tag{10}$$

Next note that (8) states that

$$\| f(x) - F(x) \| \leqq \omega(\delta). \tag{11}$$

Adding (10) and (11) gives

$$\| f(x) - \sum C_\lambda x^\lambda \| \leqq \omega(\delta) + [\pi\omega(\delta)/2\delta] \epsilon_\Lambda, \tag{12}$$

and Theorem 1 follows by choosing $\delta = \epsilon_\Lambda$.

YESHIVA UNIVERSITY.

REFERENCES.

[1] D. Jackson, "Theory of polynomial approximations," *American Mathematical Society Colloquium Publications.*

[2] Paley and Wiener, "Fourier transforms in the complex domain," *American Mathematical Society Colloquium Publications*, vol. XIX.

REAL TWO-DIMENSIONAL REPRESENTATIONS OF THE MODULAR GROUP AND RELATED GROUPS.

By J. LEHNER * and M. NEWMAN.

§ 1. Let Λ be the group of all real non-singular 2×2 matrices, and let Ω be the group of all real 2×2 matrices with determinant 1 in which a matrix is identified with its negative.[1] The problem we consider here is the determination of all faithful representations of the abstract modular group Δ (the free product of a cyclic group of order 2 and a cyclic group of order 3) by a discrete subgroup of Ω. The set of representations will be partitioned into conjugacy classes with respect to Δ in terms of a parameter ρ which assumes all values $\geqq 1$. We shall find that $\rho = 1$ gives the classical modular group

$$\Gamma = \{T, U\} = \{T\} * \{U\}, \quad T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

and that the conjugacy class determined by $\rho = 1$ includes all representations of Δ by horocyclic groups[2] and only those.

We also consider representations by subgroups of Ω which are not discrete.

Since Δ is a free product of a cyclic group of order 2 and a cyclic group of order 3, the same is true of all faithful representations of Δ. Some of our theorems give conditions under which two matrices, one of period 2, the other of period 3, generate a free product.

At the end of the paper extensions are made to the Hecke groups, which are free products of a cyclic group of order 2 and a cyclic group of order q.

The principal results of the paper are the following two theorems:

THEOREM 1. *For each* $\rho \geqq 1$ *set* $R_\rho = \begin{pmatrix} 0 & -\rho \\ 1/\rho & 1 \end{pmatrix}$, $G_\rho = \{T, R_\rho\}$. *Then*

(i) $G_\rho = \{T\} * \{R_\rho\}$, *the free product of the cyclic group* $\{T\}$ *of order* 2 *and the cyclic group* $\{R_\rho\}$ *of order* 3.

Received November 10, 1964.

* The work of this author was supported by the Office of Naval Research.

[1] If Λ_1 is the subgroup of Λ consisting of all matrices of determinant 1 and I is the identity matrix, then $\Omega = \Lambda_1/\{I, -I\}$. The group Ω is isomorphic to the group of all linear fractional transformations with real coefficients and determinant 1.

[2] A Fuchsian group is horocyclic if every real number is an accumulation point of images of a point in the upper half-plane by elements of the group.

(ii) *G_ρ is a discrete group.*

(iii) *Every discrete faithful representation of Δ by a subgroup of Ω is conjugate over Λ to G_ρ for some $\rho \geqq 1$.*

(iv) *If $\rho, \sigma \geqq 1$ then G_ρ and G_σ are conjugate over Λ if and only if $\rho = \sigma$.*

THEOREM 2. *The conjugacy class determined by $\rho = 1$ consists entirely of horocyclic groups whereas the conjugacy classes determined by ρ for all $\rho > 1$ consist entirely of non-horocyclic groups.*

§ 2. A first result is contained in the following:

THEOREM 3. *Put*

$$R = \begin{pmatrix} -a & -b \\ c & 1+a \end{pmatrix}, \quad a^2 + a + 1 = bc$$

where $a \geqq 0$, $b > 0$, $c > 0$. Then if G is the group generated by T and R, G is discrete and is the free product of the cyclic group $\{T\}$ of order 2 and the cyclic group $\{R\}$ of order 3.

Proof. We can assume that $R \neq U$, since then G is just Γ and the truth of the theorem is known in this case. It follows that $\max(b, c) > 1$.

We have that

$$TR = \begin{pmatrix} c & 1+a \\ a & b \end{pmatrix}, \quad TR^2 = \begin{pmatrix} c & a \\ 1+a & b \end{pmatrix}.$$

Thus the diagonal elements of TR, TR^2 are positive and the off-diagonal elements are non-negative. Furthermore at most one off-diagonal element can vanish in each matrix. It follows by the methods of [2] that $\{T, R\} = \{T\} * \{R\}$.

Now suppose that G is not discrete and that $\{A_n\}$, $n \geqq 1$ is an infinite sequence of distinct elements of G which converges to some limit matrix. The words of G may be divided into 4 classes: those that begin with T and end with T, begin with T and end with R, begin with R and end with T and begin with R and end with R. Therefore there must be an infinite subsequence $\{B_n\}$, $n \geqq 1$ whose elements are all in the same class; and since any class is obtainable from any other class by multiplications by T and R we can assume that the elements B_n all begin with T and end with R. Then the B_n's belong to the semigroup generated by TR, TR^2.

We have that[3]

$$TR \gg \begin{pmatrix} c & 0 \\ 0 & b \end{pmatrix}, \quad TR^2 \gg \begin{pmatrix} c & 0 \\ 0 & b \end{pmatrix}.$$

Suppose that B_n is of total length l_n in TR, TR^2. Then $l_n \to \infty$ as $n \to \infty$ and

$$B_n \gg \begin{pmatrix} c & 0 \\ 0 & b \end{pmatrix}^{l_n} = \begin{pmatrix} c^{l_n} & 0 \\ 0 & b^{l_n} \end{pmatrix}.$$

Since $\max(b, c) > 1$ it follows that the sequence $\{B_n\}$, $n \geqq 1$ cannot converge to a limit matrix. Hence G is discrete and the proof of Theorem 3 is concluded.

§ 3. We now suppose given elements A, B of Ω such that A is of period 2 and B of period 3. Then $\operatorname{tr} A = 0$, $\operatorname{tr} B = \pm 1$. Since an element of Ω is to be identified with its negative, we can assume that $\operatorname{tr} B = 1$. The group $\{A, B\}$ will be the subject of our study, and we will be interested in determining simple forms for $\{A, B\}$ by conjugacies over Λ.

LEMMA 1. *If $M \in \Omega$, then M commutes with T if and only if*

$$M = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \quad \phi \text{ real.} \tag{1}$$

We omit the easy proof.

LEMMA 2. *Let $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Omega$. Then it is possible to determine M of the form* (1) *so that MCM^{-1} has equal diagonal elements.*

Proof. By direct computation, we find that the diagonal elements of MCM^{-1} are equal if and only if

$$(a - d)\cos 2\phi = (b + c)\sin 2\phi.$$

Thus we need only choose ϕ so that

$$\tan 2\phi = \frac{a - d}{b + c},$$

a choice which clearly is always possible.

LEMMA 3. *Suppose that $\operatorname{tr} A = 0$, $A \in \Omega$. Then A is conjugate over Λ to T.*

---

[3] We write $A \gg B$ to mean that B is a matrix with non-negative entries and that the entries of A are not less than the corresponding entries of B.

Proof. We can write

$$A = \begin{pmatrix} -a & b \\ -c & a \end{pmatrix}, \quad a^2 + 1 = bc.$$

Then $c \neq 0$. Put

$$M = \begin{pmatrix} 1 & a \\ 0 & c \end{pmatrix} \in \Lambda.$$

Then it is easily verified that

$$A = MTM^{-1}.$$

LEMMA 4. *Suppose that $A, B \in \Omega$, $\operatorname{tr} A = 0$, $\operatorname{tr} B = 1$. Let $G = \{A, B\}$. Then after a suitable conjugacy by an element of Λ we can assume that $G = \{T, S\}$ where $S \in \Omega$,*

$$S = \begin{pmatrix} \frac{1}{2} & -b \\ c & \frac{1}{2} \end{pmatrix}, \quad b > 0, c > 0, bc = \tfrac{3}{4}. \tag{2}$$

Proof. By Lemma 3 A may be chosen to be T. By Lemmas 1 and 2, B may be chosen to be of the form

$$B = \begin{pmatrix} \frac{1}{2} & -b \\ c & \frac{1}{2} \end{pmatrix}, \quad bc = \tfrac{3}{4}$$

while preserving T. Then b and c are of the same sign. If b and c are positive, choose $S = B$. If b and c are negative, choose $S = B^{-1}$. This completes the proof of the lemma.

§ 4. We now turn to the proof of Theorem 1 and consider two cases, depending upon whether $b + c < 2$ or $b + c \geqq 2$. We note that

$$TS = \begin{pmatrix} c & \frac{1}{2} \\ -\frac{1}{2} & b \end{pmatrix}, \quad \operatorname{tr}(TS) = b + c.$$

We set $G = \{T, S\}$.

Suppose first that $b + c < 2$. We prove

THEOREM 4. *Suppose that $b + c < 2$. Then G cannot be simultaneously discrete and equal to the free product of a cyclic group of order 2 and a cyclic group of order 3.*

Proof. Since $b + c < 2$, TS is elliptic. If TS is of infinite order then G is not discrete. Suppose that TS is of finite order n. If G is to be the free product of a cyclic group of order 2 and a cyclic group of order 3, n must be 2 or 3; and it follows that $\operatorname{tr}(TS) = 0$ or 1. Neither of these is possible however since $\operatorname{tr}(TS) = b + c$ and $b > 0$, $c > 0$, $bc = \frac{3}{4}$. Hence the theorem is proved.

Now suppose that $b+c \geqq 2$. Then

$$(b-c)^2 = (b+c)^2 - 4bc \geqq 1,$$

since $b+c \geqq 2$ and $bc = \frac{3}{4}$. Hence

$$(3) \qquad |b-c| \geqq 1.$$

Let M be a matrix of the form (1). Then M commutes with T and the (1, 1) element of MRM^{-1} is

$$\tfrac{1}{2} + \tfrac{1}{2}(b-c)\sin 2\phi.$$

By (3), ϕ can be determined so that this element is 0. Hence after a suitable conjugacy we may assume that $G = \{T, R_\rho\}$, where

$$R_\rho = \begin{pmatrix} 0 & -\rho \\ 1/\rho & 1 \end{pmatrix}.$$

Furthermore we can assume that $\rho \geqq 1$, since the conjugacy determined by $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ carries T into $-T$ and R_ρ into $R_{-\rho}$; and the conjugacy determined by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ carries T into $-T$ and R_ρ^{-1} into $R_{1/\rho}$. Then Theorem 3 implies that $G = \{T\} * \{R_\rho\}$ and is discrete.

To complete the proof of Theorem 1 we need only show that two groups $G_\rho = \{T\} * \{R_\rho\}$ and $G_\sigma = \{T\} * \{R_\sigma\}$, $\rho, \sigma \geqq 1$ are conjugate over Λ if and only if they are identical. Suppose the groups conjugate over Λ, and suppose that $\rho \leqq \sigma$. Then every trace occurring in G_ρ must also occur in G_σ. In particular, $\operatorname{tr}(TR_\rho) = \rho + \dfrac{1}{\rho}$ must occur in G_σ. It is not difficult to show that every element of G_σ not conjugate over G_σ to a power of T or of R_σ is conjugate over G_σ to some element of the semigroup generated by TR_σ, TR_σ^2. Thus $\rho + \dfrac{1}{\rho}$ must occur as the trace of some element of this semigroup.

We have

$$TR_\sigma = \begin{pmatrix} 1/\sigma & 1 \\ 0 & \sigma \end{pmatrix}, \quad TR_\sigma^2 = \begin{pmatrix} 1/\sigma & 0 \\ 1 & \sigma \end{pmatrix}.$$

Hence

$$TR_\sigma \gg \begin{pmatrix} 1/\sigma & 0 \\ 0 & \sigma \end{pmatrix}, \quad TR_\sigma^2 \gg \begin{pmatrix} 1/\sigma & 0 \\ 0 & \sigma \end{pmatrix}.$$

It follows that if W is any word of the semigroup of total length l in TR_σ, TR_σ^2 then $\operatorname{tr} W \geqq \sigma^l + \frac{1}{\sigma^l}$. But $x + \frac{1}{x}$ is a monotone increasing function of x for $x \geqq 1$, and $\rho \leqq \sigma$. Hence ρ and σ must be equal, and the proof of Theorem 1 is concluded.

§ 5. We now give the proof of Theorem 2. Since G_1 is the modular group Γ and Γ is horocyclic, the conjugacy class determined by $\rho = 1$ contains only horocyclic groups. Suppose now that $\rho > 1$. By Lemma 4 G_ρ is conjugate over Λ to a group $G = \{T\} * \{S\}$ where

$$S = \begin{pmatrix} \frac{1}{2} & -b \\ c & \frac{1}{2} \end{pmatrix},\ b > 0,\ c > 0,\ b + c > 2,\ bc = \tfrac{3}{4}.$$

(The case $b + c = 2$ corresponds to $G_1 = \Gamma$.) We may also assume $c \geqq b$, since S may be replaced by $TST^{-1} = \begin{pmatrix} \frac{1}{2} & -c \\ b & \frac{1}{2} \end{pmatrix}$. The function

$$b + c = c + \frac{3}{4c}$$

is monotone increasing for $c \geqq \frac{\sqrt{3}}{2}$; and since $c > 1$ and $b + c > 2$, it follows that $c > \frac{3}{2}$. Consider now the isometric circles of T, S and S^{-1}, drawn below. The region F_T is a fundamental region for the group $\{T\}$, while F_S is a fundamental region for the group $\{S\}$. Since F_T contains the exterior of F_S and F_S contains the exterior of F_T, it follows that G, the free product of $\{T\}$ and $\{S\}$, has the fundamental region $F_T \cap F_S = F_G$, as shown below (see [1, p. 119]). Hence G is not horocyclic and the proof of the theorem is concluded.

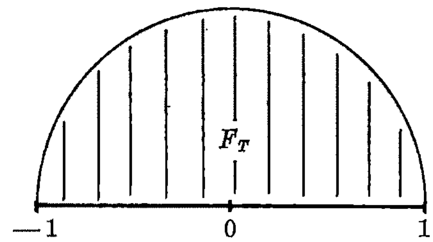
§ 6. We now turn to the case $b + c < 2$ and drop the requirement of discreteness. Thus we are considering the groups $G = \{T, S\}$ where S is given by (2) and $b + c < 2$. We note that G is *not* the free product of a cyclic group of order 2 and a cyclic group of order 3 in infinitely many cases: for example all those in which $b + c = 2 \cos \pi/n$, $n > 3$. (In these cases TS is of finite order $n > 3$.) However we can prove

THEOREM 4. *Suppose that c is transcendental. Then $G = \{T\} * \{S\}$.*

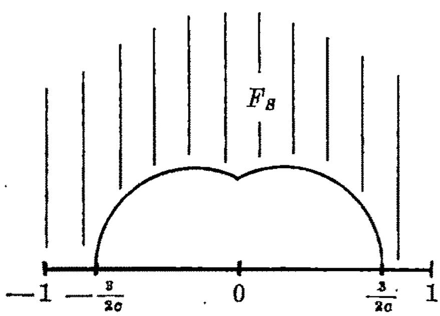
Proof. We have

$$TS = \begin{pmatrix} c & \frac{1}{2} \\ -\frac{1}{2} & 3/4c \end{pmatrix}, \qquad TS^2 = \begin{pmatrix} c & -\frac{1}{2} \\ \frac{1}{2} & 3/4c \end{pmatrix}.$$

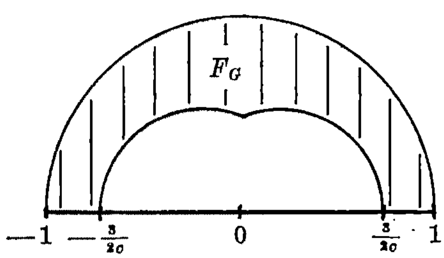
The entries of TS, TS^2 are rational functions of c with rational coefficients and the $(1,1)$ element in each case is of degree 1 in c with leading coefficient 1. Let W be any word of the semigroup generated by TS, TS^2 and suppose that W is of total length l in TS, TS^2. Then the elements of W are rational



T



S, S^{-1}



functions of c with rational coefficients, and it is easily proved by induction that the $(1,1)$ element of W is of degree l with leading coefficient 1. Hence W can never reduce to the identity and this is sufficient to prove the theorem.

An interesting corollary of Theorem 3 is the following:

THEOREM 5. *Let A, B be any elements of the classical modular group Γ such that A is of period 2 and B of period 3. Then $\{A,B\}=\{A\}*\{B\}$ and is discrete.*

Proof. That $\{A,B\}$ is discrete is clear since Γ is discrete. That $\{A,B\}=\{A\}*\{B\}$ follows from the fact that A is conjugate over Γ to T and that the normal form for B specified by Theorem 3 can always be achieved since the entries of B are integers. The operations required are replaceing B and B^{-1}, and performing the conjugacy defined by $\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}$.

§ 7. We consider certain generalizations in this section and outline the proofs of these generalizations.

There is an immediate generalization of Theorem 3 to the case of the free product of any pair of finite cyclic groups. Define

$$\lambda_n=2\cos\pi/n,\quad n\geqq 2$$

and put

$$X=\begin{pmatrix}0 & 1\\ -1 & \lambda_p\end{pmatrix},$$

$$Y=\begin{pmatrix}-a & -b\\ c & \lambda_q+a\end{pmatrix},\ a^2+\lambda_q a+1=bc.$$

Here p, q are integers $\geqq 2$ (excluding $p=q=2$) and $a\geqq 0$, $b>0$, $c>0$. Then we have

THEOREM 6. *The group $G=\{X,Y\}$ is the free product $\{X\}*\{Y\}$ of the cyclic group $\{X\}$ of order p and the cyclic group $\{Y\}$ of order q, and is discrete.*

The proof rests on the observation that the entries of

$$X^rY^s,\quad 1\leqq r\leqq p-1,\quad 1\leqq s\leqq q-1,$$

are non-negative, that the diagonal elements are positive and that at most one off-diagonal element can vanish. The omitted case $p=q=2$ can be restored by imposing the additional requirement that $XY\neq I$.

An especially important class of groups is given by

$$H_q=\{T,U_q\},$$

where

$$U_q=\begin{pmatrix}0 & -1\\ 1 & \lambda_q\end{pmatrix},\quad q\geqq 3.$$

These are the *Hecke groups*. The group H_q is discrete and is the free product $\{T\} * \{U_q\}$ of the group $\{T\}$ of order 2 and the cyclic group $\{U_q\}$ of order q. Suppose that G is any representation of H_q by a subgroup of Ω. Then there are elements A, B of Ω such that A is of period 2, B of period q and $G = \{A, B\} = \{A\} * \{B\}$. We can assume that $A = T$ by applying a suitable conjugacy over Λ to G. In order for B to be of exponent q the eigenvalues of B must be $\exp(i\pi r/q)$, $\exp(-i\pi r/q)$ for some integer r such that $(r, q) = 1$. Determine r' so that $rr' \equiv 1 \pmod{q}$. Then $(r', q) = 1$, $G = \{T\} * \{B^{r'}\}$, and the eigenvalues of $B^{r'}$ are $\exp(i\pi/q)$, $\exp(-i\pi/q)$ or their negatives. Since $B^{r'}$ can also be replaced by $-B^{r'}$, we conclude that $G = \{T\} * \{S\}$ where $\operatorname{tr} S = \lambda_q$.

By Lemma 2 a conjugacy over Λ may be applied to G which leaves T unchanged and makes the diagonal elements of S equal. It follows that S may be taken of the form

$$S = \begin{pmatrix} \cos \pi/q & -b \\ c & \cos \pi/q \end{pmatrix}, \quad bc = \sin^2 \pi/q;$$

and b and c may be chosen > 0, since S may be replaced by S^{-1}. We note that $\operatorname{tr}(TS) = b + c$.

Again the discussion falls into two cases: $b + c < 2$, $b + c \geqq 2$. In the former we can argue as before that G cannot be discrete if it is the free product $\{T\} * \{S\}$. In the latter we find that a suitable conjugacy over Λ may be applied to G which fixes T and which takes S into

$$S_\rho = \begin{pmatrix} 0 & -\rho \\ 1/\rho & \lambda_q \end{pmatrix}.$$

As before we may assume $\rho \geqq 1$. Then proceeding as in the proof of Theorem 1 we obtain

THEOREM 7. *For each* $\rho \geqq 1$ *set* $S_\rho = \begin{pmatrix} 0 & -\rho \\ 1/\rho & \lambda_q \end{pmatrix}$, $G_\rho = \{T, S_\rho\}$. *Then*

(i) $G_\rho = \{T\} * \{S_\rho\}$, *the free product of the cyclic group* $\{T\}$ *of order* 2 *and the cyclic group* $\{S_\rho\}$ *of order* q.

(ii) G_ρ *is a discrete group.*

(iii) *Every discrete representation of the Hecke group* H_q *by a subgroup of* Ω *is conjugate over* Λ *to* G_ρ *for some* $\rho \geqq 1$.

(iv) *If* $\rho, \sigma \geqq 1$ *then* G_ρ *and* G_σ *are conjugate over* Λ *if and only if* $\rho = \sigma$.

In just the way that Theorem 2 was proved we can also prove

THEOREM 8. *The conjugacy class determined by* $\rho = 1$ *for the groups defined in Theorem 7 above consists entirely of horocyclic groups, whereas the conjugacy classes determined by* ρ *for all* $\rho > 1$ *consist entirely of non-horocyclic groups.*

UNIVERSITY OF MARYLAND, MARYLAND
AND
NATIONAL BUREAU OF STANDARDS, WASHINGTON.

REFERENCES.

[1] J. Lehner, "Discontinuous groups and automorphic functions," *Mathematics surveys* 8, *American Mathematical Society* (1964).

[2] M. Newman, "Some free products of cyclic groups," *Michigan Mathematical Journal*, vol. 9 (1962), pp. 369-373.

MAXIMAL ABELIAN SUBALGEBRAS IN HYPERFINITE FACTORS.

By SALVATORE ANASTASIO.[1]

1. **Introduction.** In his seminal paper [2], Dixmier has defined three types of maximal abelian subalgebras $\mathfrak{A}^0$ in a factor $\mathfrak{A}$ as follows. Let $S(\mathfrak{A}^0)$ denote the *-algebra of all operators of the form $\sum_{i=1}^{n} \lambda_i U_i$, where the operators U_i are unitary operators in $\mathfrak{A}$ having the property that $U_i \mathfrak{A}^0 U_i^* = \mathfrak{A}^0$ and where the λ_i are scalars. Let $R(\mathfrak{A}^0)$ denote the weak closure of $S(\mathfrak{A}^0)$. Then $R(\mathfrak{A}^0)$ is a subalgebra of $\mathfrak{A}$ and $\mathfrak{A}^0 \subseteq R(\mathfrak{A}^0) \subseteq \mathfrak{A}$.

Definition 1. $\mathfrak{A}^0$ is called

(i) *Singular* if $R(\mathfrak{A}^0) = \mathfrak{A}^0$

(ii) *Regular* if $R(\mathfrak{A}^0) = \mathfrak{A}$

(iii) *Semi-regular* if $R(\mathfrak{A}^0)$ is a factor.

Using group algebras to construct factors, Dixmier showed the existence of a subalgebra of each of these types in a hyperfinite factor $\mathfrak{A}$. Pukánszky [5] later showed the existence of a denumerable infinity of *singular* subalgebras in $\mathfrak{A}$ which cannot be pairwise connected by *-automorphisms of $\mathfrak{A}$. Recently, Tauer [7] established the same result for *semi-regular* subalgebras. As for *regular* subalgebras, it may well be that all of them are conjugate under *-automorphisms of $\mathfrak{A}$, but to date this conjecture has not been proved.

In this paper we investigate a fourth alternative, not considered by Dixmier, namely the case where $R(\mathfrak{A}^0)$ is larger than $\mathfrak{A}^0$ but is not a factor. To define this case more precisely, we first extend the notion "$R(\mathfrak{A}^0)$" as follows. Let $\mathfrak{A}_1$ be an arbitrary subalgebra of $\mathfrak{A}$. Define $R^1(\mathfrak{A}_1)$ to be $R(\mathfrak{A}_1)$ and, inductively, $R^{j+1}(\mathfrak{A}_1)$ to be $R(R^j(\mathfrak{A}_1))$ $(j = 1, 2, \cdots)$. For convenience we also put $R^0(\mathfrak{A}_1) = \mathfrak{A}_1$. A new series of types of maximal abelian subalgebras can then be added to those already contained in Definition 1:

Definition 2. A maximal abelian subalgebra $\mathfrak{A}^0$ is called *m-semi-regular* $(m = 1, 2, \cdots)$ if

Received January 7, 1965.

[1] The results presented here are taken from the author's Ph. D. thesis, written under the supervision of Professor J. T. Schwartz at New York University.

(i) $R^j(\mathfrak{A}^0)$ is not a factor for $j < m$

but (ii) $R^m(\mathfrak{A}^0)$ is a factor.

(In terms of this definition, a *semi-regular* subalgebra is 1-*semi-regular*.) We may now state the main results of this paper.

THEOREM I. *In a hyperfinite factor $\mathfrak{A}$ there exists a denumerable infinity of '2-semi-regular' maximal abelian subalgebras which cannot be pairwise connected by *-automorphisms of $\mathfrak{A}$.*

THEOREM II. *In a hyperfinite factor $\mathfrak{A}$ there exists a denumerable infinity of '3-semi-regular' maximal abelian subalgebras which cannot be pairwise connected by *-automorphisms of $\mathfrak{A}$.*

(The obvious generalization of the above theorems to *m-semi-regular* maximal abelian subalgebras for every $m > 3$ is conjectured but not here proved; for $m = 1$ the case has been settled, as noted, but by methods very different from those employed here.)

In order to establish that the subalgebras in question are pairwise non-conjugate, we shall make use of the following definition:

Definition 3. (Tauer [7]). Let $\mathfrak{A}_1$ be a subalgebra of $\mathfrak{A}$. $\mathfrak{A}_1$ is said to be of *length* L (in $\mathfrak{A}$) if there is a chain[2]

$$\mathfrak{A}_1 \subset R(\mathfrak{A}_1) \subset R^2(\mathfrak{A}_1) \subset \cdots \subset R^L(\mathfrak{A}_1) = \mathfrak{A} \tag{1}$$

2. Construction of factors. To construct the desired factors, we follow von Neumann's procedure in [4]:

G shall be a countable discrete group with identity e. $\mathfrak{H}$ shall be $L_2(G)$, the Hilbert space of square-summable complex valued functions on G. For each $g \in G$ there is a unitary operator U_g defined on $\mathfrak{H}$ by $U_g x(g') = x(g'g)$. These operators generate a von Neumann algebra $\mathfrak{A}$ which is a factor of type II_1 if all the non-trivial classes of G are infinite, and which is, in addition, *hyperfinite* if G is the union of an increasing sequence of finite subgroups.

Let $\mathfrak{H}'$ be the set of those functions $y \in \mathfrak{H}$ possessing the following property: for every $x \in \mathfrak{H}$ the convolution product $x * y$ belongs to $\mathfrak{H}$. With each $y \in \mathfrak{H}'$ we associate the operator U_y defined by $U_y x = x * y$.

[2] Here and throughout the paper, the symbol $\subset$ denotes *proper* containment.

Then $\mathcal{A} = \{U_y \mid y \in \mathcal{H}'\}$ and we have:

(i) $U_y^* = U_{\tilde{y}}$ $\qquad (\tilde{y}(g) = \overline{y(g^{-1})})$

(ii) $U_{y*z} = U_z U_y$ $\qquad y, z \in \mathcal{H}'$

(iii) $U_{y+z} = U_y + U_z$ $\qquad y, z \in \mathcal{H}'$

(iv) $U_{g^{-1}} = U_{\epsilon_g}$ $\qquad$ where $\epsilon_g \in \mathcal{H}'$ is that function equal to 1 on g and 0 elsewhere.

Finally, if G_1 is a subgroup of G, the operators U_{g_1}, $g_1 \in G_1$, generate a subalgebra $\mathcal{A}_1 \subseteq \mathcal{A}$ and $\mathcal{A}_1 = \{U_z \mid z \in \mathcal{H}', z(g) = 0 \text{ if } g \notin G_1\}$.

$\mathcal{A}_1$ is called the algebra *associated with* G_1 and is abelian if and only if G_1 is abelian.

The following three lemmas are needed throughout the paper:

LEMMA 1. *Let G, $\mathcal{A}$, G_1, and $\mathcal{A}_1$ be as immediately above. Let G_1 be abelian. $\mathcal{A}_1$ is maximal abelian if and only if G_1 has the following property: (α) For every $g \in G \backslash G_1$ the set $C_g = \{g_1 g g_1^{-1} \mid g_1 \in G_1\}$ is infinite.*

LEMMA 2. *Let G and $\mathcal{A}$ be as in Lemma 1. Let G_2 be a subgroup of G and G_1 a normal subgroup of G_2. Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be the subalgebras of $\mathcal{A}$ associated with G_1 and G_2, respectively. Suppose the following property is satisfied:*

(β) *For every finite subset $B \subseteq G$ and every $g_0 \in G \backslash G_2$ there exists $g_1 \in G_1$ such that* (i) $g_0 g_1 g_0^{-1} \notin G_1$ *and* (ii) $g, g' \in B$ *and* $g' g_1 g^{-1} = g_1$ *together imply* $g = g'$.

Then $R(\mathcal{A}_1) = \mathcal{A}_2$.

Lemmas 1 and 2 are due to Dixmier [2] except that in Lemma 2 we have removed Dixmier's hypothesis that G_1 be abelian. The proof, which is independent of the hypothesis, is omitted.

LEMMA 3. *Let G and $\mathcal{A}$ be as in the previous two lemmas. Let G_1 be a subgroup of G and $\mathcal{A}_1$ its associated subalgebra. Let Z be the center of G_1 and $\mathcal{A}_Z$ its associated subalgebra. Let $\mathcal{Z} = \mathcal{A}_1 \cap \mathcal{A}_1'$ be the center of $\mathcal{A}_1$. If Z has the property that the set $D_g = \{g_1 g g_1^{-1} \mid g_1 \in G_1\}$ is infinite for each $g \in G_1 \backslash Z$, then $\mathcal{Z} = \mathcal{A}_Z$.*

Proof. (a) Clearly $\mathcal{A}_Z \subseteq \mathcal{Z}$ since $\mathcal{A}_Z$ is generated by the operators U_g, $g \in Z$, and $U_{g_1} U_g U_{g_1^{-1}} = U_{g_1 g g_1^{-1}} = U_g$ for all $g_1 \in G_1$.

(b) Let $A \in \mathfrak{F}$. Then $A = U_y$ where y vanishes outside of G_1. We must show that y vanishes outside of Z. If U_g is a generator of $\mathfrak{A}_1$ (i.e. $g \in G_1$) we have

$$A = U_y = U_{g^{-1}} U_y U_g = U_{\epsilon_g} U_y U_{\epsilon_g^{-1}} = U_{\epsilon_g^{-1} * y * \epsilon_g}.$$

But

$$(\epsilon_{g^{-1}} * y * \epsilon_g)(g') = \sum_{h \in G_1} (\epsilon_{g^{-1}} * y)(h) \epsilon_g(h^{-1} g') = (\epsilon_{g^{-1}} * y)(g' g^{-1}) = y(g g' g^{-1}).$$

Therefore, $y(g') = y(gg'g^{-1})$ for all g', $g \in G_1$. But if $g' \notin Z$, the class $D_{g'}$ is infinite. Since $\| y \|_2 < \infty$, $y(g') = 0$. That is, y vanishes outside of Z. Q. E. D.

The final lemma of this section establishes the fact that the length of a subalgebra is a *-algebraic invariant.

Lemma 4. *Let σ be a *-automorphism of a von Neumann algebra $\mathfrak{A}$ and let $\mathfrak{A}_1$ and $\mathfrak{A}_2$ be subalgebras of $\mathfrak{A}$ of lengths L_1 and L_2, respectively. If $\sigma(\mathfrak{A}_1) = \mathfrak{A}_2$, then $L_1 = L_2$.*

Proof. (a) We first show that $\sigma(R(\mathfrak{A}_1)) = R(\mathfrak{A}_2)$. Let U_1 be any unitary operator in $\mathfrak{A}$ for which $U_1 \mathfrak{A}_1 U_1^* = \mathfrak{A}_1$. Then

$$\mathfrak{A}_2 = \sigma(\mathfrak{A}_1) = \sigma(U_1 \mathfrak{A}_1 U_1^*) = \sigma(U_1) \mathfrak{A}_2 [\sigma(U_1)]^*.$$

Therefore, $\sigma(U_1)$ is one of the generators of the algebra $S(\mathfrak{A}_2)$. Conversely, every generator U_2 of $S(\mathfrak{A}_2)$ is of this form. Therefore, $\sigma(S(\mathfrak{A}_1)) = S(\mathfrak{A}_2)$.

Now, let $A \in R(\mathfrak{A}_1)$. Then A is the weak limit of a sequence of operators A_n in $S(\mathfrak{A}_1)$. Since the notion of weak convergence is "purely algebraic" (i.e. can be expressed in terms of the identity operator and the operations λA, A^*, $A + B$, and AB [von Neumann, 8]) it follows that $\sigma(A)$ is the weak limit of the sequence $\sigma(A_n)$ in $S(\mathfrak{A}_2)$ and is consequently in $R(\mathfrak{A}_2)$. Conversely, if $B \in R(\mathfrak{A}_2)$ then $\sigma^{-1}(B) \in R(\mathfrak{A}_1)$. Therefore, $\sigma(R(\mathfrak{A}_1)) = R(\mathfrak{A}_2)$.

(b) Assume that $\sigma(R^{k-1}(\mathfrak{A}_1)) = R^{k-1}(\mathfrak{A}_2)$.
Then

$$\begin{aligned} \sigma(R^k(\mathfrak{A}_1)) &= \sigma(R(R^{k-1}(\mathfrak{A}_1))) \\ &= R(\sigma(R^{k-1}(\mathfrak{A}_1))) = R^k(\mathfrak{A}_2). \end{aligned}$$

Therefore, supposing $L_1 \leqq L_2$, we have

$$\sigma(R^k(\mathfrak{A}_1)) = R^k(\mathfrak{A}_2) \qquad 1 \leqq k \leqq L_1$$

But then

$$\mathfrak{A} = \sigma(\mathfrak{A}) = \sigma(R^{L_1}(\mathfrak{A}_1)) = R^{L_1}(\mathfrak{A}_2)$$

and $L_1 = L_2$ by the definition of length.

3. Proof of Theorem I. Here and henceforth, let F denote an infinite commutative field which is the union of an increasing sequence of finite subfields. Then $F = \bigcup_{i=1}^{\infty} F_i$ where F_i are finite fields and $F_1 \subseteq F_2 \subseteq \cdots$.

We construct an infinite sequence of subalgebras as follows. For each $n \geqq 4$, we construct a group G_n of $n \times n$ matrices over F and an abelian subgroup G_{n0} such that

$$(2) \qquad G_{n0} \subset N_{n1} \subset N_{n2} \subset \cdots \subset N_{n,n-2} = G_n$$

where N_{n1} is the normalizer of G_{n0} and, for each k $(2 \leqq k \leqq n-2)$, N_{nk} is the normalizer of $N_{n,k-1}$.

If $\mathcal{A}_n$ denotes the algebra associated with G_n, $\mathcal{A}_{n0}$ the algebra associated with G_{n0} and, for each k $(1 \leqq k \leqq n-2)$, $\mathcal{A}_{nk}$ the algebra associated with N_{nk}, we then have

$$(3) \qquad \mathcal{A}_{n0} \subset \mathcal{A}_{n1} \subset \cdots \subset \mathcal{A}_{n,n-2} = \mathcal{A}_n.$$

After showing that $\mathcal{A}_n$ is a hyperfinite factor and that $\mathcal{A}_{n0}$ is maximal abelian, Lemma 2 will be used to show that $R(\mathcal{A}_{n0}) = \mathcal{A}_{n1}$ and $R(\mathcal{A}_{n,k-1}) = \mathcal{A}_{nk}$, $2 \leqq k \leqq n-2$. It will therefore follow that

$$(4) \qquad \mathcal{A}_{n0} \subset R(\mathcal{A}_{n0}) \subset R^2(\mathcal{A}_{n0}) \subset \cdots \subset R^{n-2}(\mathcal{A}_{n0}) = \mathcal{A}_n$$

and, consequently, that for each $n \geqq 4$, $\mathcal{A}_{n0}$ is of length $n-2$. We then establish, finally, that $R(\mathcal{A}_{n0})$ is not a factor but $R^2(\mathcal{A}_{n0})$ is a factor so that $\mathcal{A}_{n0}$ is 2-semi-regular.

To get down to particulars, G_n shall be the group of $n \times n$ $(n \geqq 4)$ matrices[3] of the form

$$(5) \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & 1 & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & 1 & \cdots & a_{3,n-1} & a_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

where all the entries are taken from the field F and $a_{11} \neq 0$. $\mathcal{A}_n$ shall be the von Neumann algebra generated by the operators U_g, $g \in G_n$. Lemma 5 establishes that $\mathcal{A}_n$ is a factor of type II_1.

LEMMA 5. *The class $A_{\bar{g}} = \{g\bar{g}g^{-1} \mid g \in G_n\}$ is infinite for each $\bar{g} \in G_n$, $\bar{g} \neq e$.*

---

[3] Straight bars are used throughout for all matrices.

Proof. For $n=4$, let g be of the form (5) and let $\bar{g}$ be of the same form but with entries b_{ij}. Then $g\bar{g}g^{-1}$ is the matrix

$$\begin{vmatrix} b_{11} & a_{12}(1-b_{11})+a_{11}b_{12} & X & Y \\ 0 & 1 & b_{23} & b_{24}+a_{13}b_{34}-b_{23}a_{34} \\ 0 & 0 & 1 & b_{34} \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

where $X=b_{11}(a_{12}a_{23}-a_{13})-a_{11}b_{12}a_{13}-a_{12}a_{23}+a_{11}b_{13}+a_{12}b_{23}+a_{13}$
and $Y=b_{11}(a_{12}a_{24}+a_{13}a_{34}-a_{14}-a_{12}a_{23}a_{34})+(a_{11}b_{12}+a_{12})(a_{23}a_{34}-a_{24})$
$\qquad -a_{34}(a_{11}b_{13}+a_{12}b_{23}+a_{13})+a_{11}b_{14}+a_{12}b_{24}+a_{13}b_{34}+a_{14}$.

Consider $(1,2)$ (i.e., the entry in the first row, second column). If either $b_{11}\neq 1$ or $b_{12}\neq 0$, varying a_{12} or a_{11} will produce infinitely many matrices. If neither, consider $(2,4)$. If either b_{34} or b_{23} is non-zero we can again produce infinitely many matrices. If $b_{34}=b_{23}=0$, consider $(1,3)=X=a_{11}b_{13}$. If $b_{13}\neq 0$ we have infinitely many matrices. If $b_{13}=0$, then $Y=a_{11}b_{14}+a_{12}b_{24}$ and, since $\bar{g}\neq e$, not both b_{14} and b_{24} are zero, and the lemma holds for $n=4$.

Assume the lemma true for $n-1\geqq 4$. Let

$$\hat{A}_{\bar{g}}=\{g\bar{g}g^{-1}\mid g\in G_{n-1}\},\ \bar{g}\in G_{n-1},\ \bar{g}\neq e.$$

Then $\hat{A}_{\bar{g}}$ is infinite by induction hypothessis.

Let $\tilde{A}_{\bar{g}}=\{g\bar{g}g^{-1}\mid g\in S_n\}$, $\bar{g}\in G_n$, $\bar{g}\neq e$, where $S_n\subset G_n$ is that set of elements for which all the entries in the n-th column are zero (except, of course, $a_{nn}=1$).

Since $\tilde{A}_{\bar{g}}\subseteq A_{\bar{g}}$ for each $\bar{g}$, it is enough to show that $\tilde{A}_{\bar{g}}$ is infinite for each $\bar{g}\neq e$.

To this end, let $g\in S_n$ be partitioned as follows:

$$\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1,n-1} & 0 \\ 0 & 1 & \cdots & b_{2,n-1} & 0 \\ 0 & 0 & \cdots & b_{3,n-1} & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 0 \\ \hline 0 & 0 & \cdots & 0 & 1 \end{vmatrix} = \begin{vmatrix} g_{11} & 0 \\ \hline 0 & 1 \end{vmatrix} \tag{6}$$

Let $\bar{g}$ be the matrix (5) but partitioned as follows:

$$\bar{g}=\begin{vmatrix} \bar{g}_{11} & \bar{g}_{12} \\ \hline 0 & 1 \end{vmatrix}$$

where $\bar{g}_{11}$ is an $(n-1)\times(n-1)$ matrix.

Then $g\bar{g}g^{-1}$ is the matrix

$$\left|\begin{array}{c|c} g_{11}\bar{g}_{11}g_{11}^{-1} & g_{11}\bar{g}_{12} \\ \hline 0 & 1 \end{array}\right|$$

If $\bar{g}_{11}$ is not the identity of G_{n-1}, $\{g_{11}\bar{g}_{11}g_{11}^{-1}\}$ is infinite and so is $\tilde{A}_{\bar{g}}$. If $\bar{g}_{11}=e$ then one of $a_{1n}, a_{2n}, \cdots, a_{n-1,n}$ is not zero. Say $a_{jn}\neq 0$ $(1\leqq j\leqq n-1)$. The entry in the first row of the $(n-1)\times 1$ matrix $g_{11}\bar{g}_{12}$ is $b_{11}a_{1n}+b_{12}a_{2n}+\cdots+b_{1,n-1}a_{n-1,n}$. By varying b_{1j} we can produce infinitely many matrices $g_{11}\bar{g}_{12}$ and, consequently, $\tilde{A}_{\bar{g}}$ is infinite. Q. E. D.

Since $G_n=\bigcup_{i=1}^{\infty} G_{ni}$ where G_{ni} is the *finite* group of matrices of the same form as G_n but with entries restricted to the finite field F_i, G_n is the union of an increasing sequence of finite subgroups and $\mathfrak{A}_n$ is hyperfinite.

For each $n\geqq 4$, let G_{n0} be the abelian subgroup of G_n whose typical element is

$$(7)\qquad g_0=\left|\begin{array}{cccccccc} 1 & a_{12} & a_{13} & a_{14} & a_{15} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & 1 & a_{12} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{array}\right|$$

If $\mathfrak{A}_{n0}$ is the subalgebra of $\mathfrak{A}_n$ associated with G_{n0} then

LEMMA 6. *$\mathfrak{A}_{n0}$ is a maximal abelian subalgebra of $\mathfrak{A}_n$.*

Proof. Using Lemma 1 we need only establish that G_{n0} possesses the property (α). The proof is again by induction, based on the case $n=4$. Here, let $\bar{g}\in G_4\backslash G_{40}$ be of form (5) but with the entries b_{ij} and g_0 of form (7). Since $\bar{g}\notin G_{40}$, at least one of the following must hold:

(i) $b_{11}\neq 1$ (iii) $b_{24}\neq 0$

(ii) $b_{12}\neq b_{23}$ (iv) $b_{34}\neq 0$.

Consider the matrix $g_0\bar{g}g_0^{-1}$:

$(1,2)=a_{12}(1-b_{11})+b_{12}$.

$(2,4)=b_{24}+a_{12}b_{34}$.

$(1,3)=a_{12}{}^2(b_{11}-1)+a_{13}(1-b_{11})+a_{12}(b_{23}-b_{12})+b_{13}$.

$(1,4)=a_{14}(1-b_{11})+b_{14}+a_{12}b_{24}+a_{13}b_{34}$.

Clearly any one of (i)-(iv) is enough to guarantee the desired result for $n=4$.

Assuming the lemma true for $n-1\geqq 4$, let $\bar{g}\in G_n\backslash G_{n0}$ be of form (5), partitioned as in Lemma 5, and $g\in G_{n0}$ be of form (7) but with the entries b_{ij}. Then, reasoning as in the previous lemma, if $\bar{g}_{11}\notin G_{n-1,0}$ we are done. If $\bar{g}_{11}\in G_{n-1,0}$ then at least one of $a_{2n}, a_{3n},\cdots, a_{n-1,n}$ is not zero. The first entry in the $(n-1)\times 1$ matrix in the upper right corner of $g\bar{g}g^{-1}$ is $a_{1n}+b_{12}a_{2n}+b_{13}a_{3n}+\cdots+b_{1,n-1}a_{n-1,n}$. The last two sentences complete the proof. Q. E. D.

It should be noted that, in Lemma 2, the fact that G_1 is normal in G_2 together with the requirement (i) of (β) entail that G_2 is in fact the *normalizer* of G_1.

LEMMA 7. *The normalizer, N_{n1}, of G_{n0} is the group of matrices of the form:*

$$(8)\qquad \begin{vmatrix} 1 & a_{12} & a_{13} & a_{14} & a_{15} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & 1 & a_{23} & a_{24} & a_{25} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & a_{45} & \cdots & a_{4,n-1} & a_{4n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

(i. e. those elements of G_n of form (5) but with $a_{3k}=0$ for $4\leqq k\leqq n$ and $a_{11}=1$.)

Proof. A straightforward computation serves to establish the result for $n=4$. Assuming the result for $n-1\geqq 4$ and partitioning the $n\times n$ matrices as in the previous two lemmas leads to the proof of Lemma 7 for general n. Q. E. D.

The next step is to show that N_{n1} has the property (β) with respect to G_{n0}. The following definition and the two succeeding lemmas accomplish this:

Definition 4. K_{n0} is the subset of G_{n0} containing all the elements of the form

$$(9)\qquad \begin{vmatrix} 1 & b_{12} & 0 & 0 & \cdots & 0 \\ 0 & 1 & b_{12} & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{vmatrix}$$

where $b_{12}\neq 0$.

LEMMA 8. *If* $g_0 \in K_{n0}$ *and* $\bar{g} \in G_n \backslash N_{n1}$, *then* $\bar{g} g_0 \bar{g}^{-1} \notin G_{n0}$.

Proof. A simple calculation establishes the result for $n = 4$. Proceeding inductively once more, let $\bar{g}$ be the matrix (5) partitioned as in Lemma 5. Let g_0, of form (9), be written as

$$g_0 = \left|\begin{array}{c|c} g_{11} & g_{12} \\ \hline 0 & 1 \end{array}\right| \text{ with } g_{11} \in K_{n-1,0}. \text{ Then}$$

$$\bar{g} g_0 \bar{g}^{-1} = \left|\begin{array}{c|c} \bar{g}_{11} g_{11} \bar{g}_{11}^{-1} & \hat{g} \\ \hline 0 & 1 \end{array}\right|$$

where $\hat{g} = -\bar{g}_{11} g_{11} \bar{g}_{11}^{-1} + \bar{g}_{11} g_{12} + \bar{g}_{12}$.

If $\bar{g} \notin N_{n1}$ then either (a) $\bar{g}_{11} \notin N_{n-1,1}$ or (b) $\bar{g}_{11} \in N_{n-1,1}$ but $a_{3n} \neq 0$.

Ad (a): Then $\bar{g}_{11} g_{11} \bar{g}_{11}^{-1} \notin G_{n-1,0}$ by induction hypothesis, since $b_{12} \neq 0$ is the same entry in both g_0 and g_{11}. But then $\bar{g} g_0 \bar{g}^{-1} \notin G_{n0}$.

Ad (b): $\hat{g}$ is an $(n-1) \times 1$ matrix with second row $-b_{12} a_{3n}$. If $a_{3n} \neq 0$, $\bar{g} g_0 \bar{g}^{-1} \notin G_{n0}$. Q. E. D.

LEMMA 9. *Let* B_n *be any finite subset of* G_n. *For all but finitely many elements of* K_{n0} *the following statement is true*: $u, v \in B_n$ *and* $u g_0 v^{-1} = g_0$ *imply* $u = v$ $(g_0 \in K_{n0})$.

Proof. (a) $n = 4$. Let $u, v \in B_4$, both of form (5), $u = |a_{ij}|$ and $v = |c_{ij}|$. Let $g_0 = |b_{ij}|$ be of form (9). If $u g_0 v^{-1}$ is the matrix $|d_{ij}|$, we have the following:

(i) $d_{11} = a_{11}/c_{11}$

(ii) $d_{23} = b_{12} + a_{23} - c_{23}$

(iii) $d_{34} = a_{34} - c_{34}$.

Hence, if $u g_0 v^{-1}$ is to be equal to g_0, we must have $a_{11} = c_{11}$, $a_{23} = c_{23}$, and $a_{34} = c_{34}$.

Further,

(iv) $d_{12} = (a_{11} b_{12} + a_{12} - a_{11} c_{12})/c_{11}$.

If $b_{12} \notin A_1 = \{(c_{12} - a_{12})/(a_{11} - 1)\}$ (where a_{11}, a_{12}, c_{12} are garnered from all $u, v \in B_4$—so that A_1 is finite) then $d_{12} \neq b_{12}$ unless $a_{11} = 1$, in which case $a_{12} = c_{12}$. Hence we may assume $a_{11} = c_{11} = 1$ and $a_{12} = c_{12}$. In similar manner, investigation of d_{24}, d_{13}, and d_{14} leads to finite sets A_2, A_3, A_4 such that, for any $g_0 = |b_{ij}| \in K_{40}$ for which $b_{12} \notin A_1 \cup A_2 \cup A_3 \cup A_4$, $u g_0 v^{-1} = g_0$ implies $u = v$, and the lemma is established for $n = 4$.

(b) Now define B_n' to be the set of "truncated" matrices formed by the elimination of the n-th row and the n-th column of all the matrices in B_n. Then B_n' is a finite subset of G_{n-1}. If u, v are in B_n and u', v' are their "truncations" then, by induction hypothesis, we may assume that for all but a finite number of elements of $K_{n-1,0}$ the following statement is true: $u', v' \in B_n'$ and $u'g_0'(v')^{-1} = g_0'$ imply $u' = v'$ $(g_0' \in K_{n-1,0})$.

For every $g_0' \in K_{n-1,0}$ put

$$g_0 = \left|\begin{array}{c|c} g_0' & 0 \\ \hline 0 & 1 \end{array}\right| \qquad (g_0, g_0' \text{ each of form (9)})$$

Then $g_0 \in K_{n0}$ and, further, every $g_0 \in K_{n0}$ is of this form.

Let $u = |a_{ij}|$ and $v = |c_{ij}|$ be partitioned as follows:

$$u = \left|\begin{array}{c|c} u' & u_{12} \\ \hline 0 & 1 \end{array}\right| \qquad \text{and } v = \left|\begin{array}{c|c} v' & v_{12} \\ \hline 0 & 1 \end{array}\right|$$

where u_{12} and v_{12} are $(n-1) \times 1$ matrices.

Then, ug_0v^{-1} is the matrix

$$\left|\begin{array}{c|c} u'g_0'(v')^{-1} & w_{12} \\ \hline 0 & 1 \end{array}\right|$$

where $w_{12} = -u'g_0'(v')^{-1}v_{12} + u_{12}$.

If $ug_0v^{-1} = g_0$, then $u'g_0'(v')^{-1} = g_0'$. Using the induction hypothesis, we may exclude a finite number of the g_0' in order to assure that $u' = v'$. Then $w_{12} = -g_0'v_{12} + u_{12}$ is the following matrix:

$$\left|\begin{array}{c} a_{1n} - c_{1n} - b_{12}c_{2n} \\ a_{2n} - c_{2n} - b_{12}c_{3n} \\ a_{3n} - c_{3n} \\ a_{4n} - c_{4n} \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1,n} - c_{n-1,n} \end{array}\right|$$

Clearly, if $ug_0v^{-1} = g_0$, then $a_{jn} = c_{jn}$, $3 \leqq j \leqq n-1$.

And, if we further exclude b_{12} from the union of the sets $\{(a_{1n} - c_{1n})/c_{2n}\}$ and $\{(a_{2n} - c_{2n})/c_{3n}\}$, we cannot have $w_{12} = 0$ unless $c_{2n} = c_{3n} = 0$ in which case $a_{1n} = c_{1n}$ and $a_{2n} = c_{2n}$. Therefore, $a_{jn} = c_{jn}$ for $1 \leqq j \leqq n-1$ so that $u_{12} = v_{12}$ and $u = v$. Q. E. D.

As a result of Lemmas 8 and 9 we know that N_{n1} has the property (β) with respect to G_{n0} so that, by Lemma 2, $R(\mathfrak{A}_{n0}) = \mathfrak{A}_{n1}$.

LEMMA 10. *The normalizer* N_{n2} *of* N_{n1} *is the group of those matrices* $g \in G_n$ *of general form* (10) (below) *but with the third row* $0\ 0\ 1\ \underbrace{0 \cdots 0}_{(n-4)\text{ zeros}}\ b_{3n}$.

Proof. For $n = 4$, $N_{n2} = G_n$.

Assuming the result to be true for $n-1 \geqq 4$, let $\bar{g} \in N_{n1}$ be of form (10) and partitioned as:

$$\bar{g} = \left|\begin{array}{c|c} \bar{g}_{11} & \bar{g}_{12} \\ \hline 0 & 1 \end{array}\right|, \qquad \bar{g}_{11} \in N_{n-1,1}.$$

To find all matrices $g \in G_n$ such that $g\bar{g}g^{-1} \in N_{n1}$ for all $\bar{g} \in N_{n1}$, let g be the matrix

$$(10) \qquad \left|\begin{array}{cccccc|c} b_{11} & b_{12} & b_{13} & \cdots & b_{1,n-1} & & b_{1n} \\ 0 & 1 & b_{23} & \cdots & b_{2,n-1} & & b_{2n} \\ \cdot & \cdot & \cdot & \cdots & \cdot & & \cdot \\ 0 & 0 & 0 & \cdots & 1 & & b_{n-1,n} \\ \hline 0 & 0 & 0 & \cdots & 0 & & 1 \end{array}\right| = \left|\begin{array}{c|c} g_{11} & g_{12} \\ \hline 0 & 1 \end{array}\right|$$

Then,

$$g\bar{g}g^{-1} = \left|\begin{array}{c|c} g_{11}\bar{g}_{11}g_{11}^{-1} & \hat{g} \\ \hline 0 & 1 \end{array}\right|$$

where $\hat{g} = -g_{11}\bar{g}_{11}g_{11}^{-1}g_{12} + g_{11}\bar{g}_{12} + g_{12}$.

If $g\bar{g}g^{-1}$ is to be in N_{n1} for all $\bar{g} \in N_{n1}$ then $g_{11}\bar{g}_{11}g_{11}^{-1}$ must be in $N_{n-1,1}$ for all $\bar{g}_{11} \in N_{n-1,1}$. But then $g_{11} \in N_{n-1,2}$ by induction hypothesis so that $b_{3j} = 0$, $4 \leqq j \leqq n-2$. Thus, we need only show that $b_{3,n-1} = 0$.

The entry in the third row of the $(n-1) \times 1$ matrix $\hat{g}$ is $b_{3,n-1}a_{n-1,n}$. If $g\bar{g}g^{-1}$ is to be in N_{n1} for *all* $\bar{g} \in N_{n1}$, then, we must have $b_{3,n-1} = 0$; i.e. $\bar{g} \in N_{n2}$.

The following corollary will prove useful in the sequel:

COROLLARY 1. *If* $\bar{N}_{n1}$ *denotes that subgroup of* G_n *with third row*: $0010\cdots 0$ (*so that* N_{n1} *is that subgroup of* $\bar{N}_{n1}$ *for which* $(1,1) = 1$) *then* N_{n2} *is the normalizer of* $\bar{N}_{n1}$.

Proof. The following calculation, along with the proof of Lemma 10, suffices to prove this corollary:

$$\begin{vmatrix} b_{11} & b_{12} \\ 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} a_{11} & a_{12} \\ 0 & 1 \end{vmatrix} \cdot \begin{vmatrix} 1/b_{11} & -b_{12}/b_{11} \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} a_{11} & c \\ 0 & 1 \end{vmatrix}$$

Definition 5. N_{nk} is the subgroup of G_n consisting of all matrices $g \in G_n$, of form (10), with third row:

$$0\,0\,1\,\underbrace{0\cdots 0}_{n-k-2 \text{ zeros}}\; b_{3,n-k+2}\, b_{3,n-k+3}\cdots b_{3n}$$

where $2 \leqq k \leqq n-2$.

LEMMA 11. *For each fixed n, the normalizer of N_{nk} is $N_{n,k+1}$* $(2 \leqq k \leqq n-2)$.

Proof. It may be worthwhile to remark that the following proof is *not* by induction.

Let $\bar{g} \in N_{nk}$ be the following matrix:

$$\left|\begin{array}{cccccc|ccc}
a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1,n-k+1} & a_{1,n-k+2} & \cdots & a_{1n} \\
0 & 1 & a_{23} & a_{24} & \cdots & a_{2,n-k+1} & a_{2,n-k+2} & \cdots & a_{2n} \\
0 & 0 & 1 & 0 & \cdots & 0 & a_{3,n-k+2} & \cdots & a_{3n} \\
0 & 0 & 0 & 1 & \cdots & a_{4,n-k+1} & a_{4,n-k+2} & \cdots & a_{4n} \\
\cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\
0 & 0 & 0 & 0 & \cdots & 1 & a_{n-k+1,n-k+2} & \cdots & a_{n-k+1,n} \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1
\end{array}\right|$$

Partition $\bar{g}$ as $\bar{g} = \left|\begin{array}{c|c} \bar{g}_{11} & \bar{g}_{12} \\ \hline 0 & \bar{g}_{22} \end{array}\right|$

and note that $\bar{g}_{11} \in \bar{N}_{n-k+1,1}$.

Let g be the matrix (10) partitioned as:

$$g = \left|\begin{array}{c|c} g_{11} & g_{12} \\ \hline 0 & 1 \end{array}\right|$$

where $g_{11} \in G_{n-k+1}$.

Then $g\bar{g}g^{-1}$ is the matrix

$$\left|\begin{array}{c|c} g_{11}\bar{g}_{11}g_{11}^{-1} & \hat{g} \\ \hline 0 & g_{22}\bar{g}_{22}g_{22}^{-1} \end{array}\right|$$

$g\bar{g}g^{-1}$ will be in N_{nk} for all $\bar{g} \in N_{nk}$ if and only if $g_{11}\bar{g}_{11}g_{11}^{-1}$ is in $\bar{N}_{n-k+1,1}$ for all $\bar{g}_{11} \in \bar{N}_{n-k+1,1}$. But then, by Corollary 1, g_{11} must be in $N_{n-k+1,2}$, which is to say that $g \in N_{n,k+1}$. Q. E. D.

Note: If $k = n-3$, $N_{n,k+1} = N_{n,n-2} = G_n$.

We have now demonstrated the existence of the promised string (2) of normalizers.

The next main step is to prove that $N_{n,k+1}$ has property (β) with respect to N_{nk}, $1 \leqq k \leqq n-4$ (Since $N_{n,n-2} = G_n$ has (β) trivially with respect to $N_{n,n-3}$, we need not consider $k = n-3$). The following definition and lemma pave the way.

Definition 6. M_{n1} is that subset of N_{n1} consisting of all the matrices of the form

$$\bar{g} = \begin{Vmatrix} 1 & a_{12} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & a_{23} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & a_{45} & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{Vmatrix}$$

where $a_{j,j+1} \neq 0$ for $j \neq 3$, $1 \leqq j \leqq n-1$.

(Note that $M_{n1} \subseteq N_{nk}$ for all k, $1 \leqq k \leqq n-2$.)

LEMMA 12. *If $\bar{g} \in M_{n1}$ and $g \in G_n \backslash N_{n2}$, then $g\bar{g}g^{-1} \notin \tilde{N}_{n1}$.*

Proof. Let $g \in G_n \backslash N_{n2}$ be the matrix (10). Then at least one of the entries $b_{34}, b_{35}, \cdots, b_{3,n-1}$ is non-zero. Let j' be the smallest index j such that $b_{3j} \neq 0$, $4 \leqq j' \leqq n-1$. Thus, the third row of g is

$$0 \quad 0 \quad 1 \quad 0 \cdots 0 \quad b_{3j'} \quad b_{3,j'+1} \cdots b_{3n}$$

and the third row of $g\bar{g}$ is

$$0 \quad 0 \quad 1 \quad 0 \cdots 0 \quad b_{3j'} \quad (b_{3,j'+1} + b_{3j'}a_{j',j'+1}) \cdots b_{3n}.$$

Comparing these two rows and making note of

(i) the $(j'+1)$-st column of g^{-1} is zero below the main diagonal and

(ii) in gg^{-1}, $(3, j'+1) = 0$,

it is clear that the entry $(3, j'+1)$ in $g\bar{g}g^{-1}$ contains the single term $b_{3j'}a_{j',j'+1} \neq 0$.

Hence, $g\bar{g}g^{-1}$ has a non-zero entry in the third row, $(j'+1)$-st column and, consequently, $g\bar{g}g^{-1} \notin \tilde{N}_{n1}$. Q. E. D.

LEMMA 13. *$N_{n,k+1}$ has property (β) with respect to N_{nk}, $1 \leqq k \leqq n-3$.*

Proof. (a) Let $g \in G_n \backslash N_{n,k+1}$ be the matrix

$$\left|\begin{array}{cccccc|ccc} b_{11} & b_{12} & b_{13} & \cdots & b_{1,n-k} & b_{1,n-k+1} & b_{1,n-k+2} & \cdots & b_{1n} \\ 0 & 1 & b_{23} & \cdots & b_{2,n-k} & b_{2,n-k+1} & b_{2,n-k+2} & \cdots & b_{2n} \\ 0 & 0 & 1 & \cdots & b_{3,n-k} & b_{3,n-k+1} & b_{3,n-k+2} & \cdots & b_{3n} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 1 & b_{n-k+1,n-k+2} & \cdots & b_{n-k+1,n} \\ \hline \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{array}\right|$$

$$= \left|\begin{array}{c|c} g_{11} & g_{12} \\ \hline 0 & g_{22} \end{array}\right|.$$

Since $g \notin N_{n,k+1}$ at least one of $b_{34}, b_{35}, \cdots, b_{3,n-k}$ is not zero. Therefore, $g_{11} \in G_{n-k+1} \backslash N_{n-k+1,2}$.

Let $\bar{g}$ be arbitrary in M_{n1},

$$\bar{g} = \left|\begin{array}{c|c} \bar{g}_{11} & \bar{g}_{12} \\ \hline 0 & \bar{g}_{22} \end{array}\right|, \ \bar{g}_{11} \in M_{n-k+1,1}.$$

Then, $g\bar{g}g^{-1} = \left|\begin{array}{c|c} g_{11}\bar{g}_{11}g_{11}^{-1} & \hat{g} \\ \hline 0 & g_{22}\bar{g}_{22}g_{22}^{-1} \end{array}\right|$

By Lemma 12, $g_{11}\bar{g}_{11}g_{11}^{-1} \notin \tilde{N}_{n-k+1,1}$. Consequently, $g\bar{g}g^{-1} \notin N_{nk}$. Therefore, condition (i) of (β) is satisfied: given $g \in G_n \backslash N_{n,k+1}$ we can find $\bar{g} \in N_{nk}$ such that $g\bar{g}g^{-1} \notin N_{nk}$.

(b) Let B_n be a finite subset of G_n. We then claim: for all but a finite number of elements of M_{n1}, $u, v \in B_n$ and $u\bar{g}v^{-1} = \bar{g}$ $(\bar{g} \in M_{n1})$ together imply $u = v$.

What is claimed here was established for $n = 4$ in Lemma 9. Assume now that it is true for $n - 1 \geqq 4$. Let $u, v \in B_n$ be the same matrices used in Lemma 9, and let B_n' be also as in that lemma. Let $\bar{g} \in M_{n1}$ be as in Definition 6, partitioned as:

$$\bar{g} = \left|\begin{array}{c|c} \bar{g}_{11} & \bar{g}_{12} \\ \hline 0 & 1 \end{array}\right|, \ \bar{g}_{11} \in M_{n-1,1},$$

where $\bar{g}_{12}$ is the $(n-1) \times 1$ matrix all of whose entries are zero except for $a_{n-1,n} \neq 0$ to be determined.

Then $u\bar{g}v^{-1}$ is the matrix

$$\left|\begin{array}{c|c} u'\bar{g}_{11}(v')^{-1} & \hat{g} \\ \hline 0 & 1 \end{array}\right|$$

where $\hat{g} = -u'\bar{g}_{11}(v')^{-1}v_{12} + u'\bar{g}_{12} + u_{12}$.

Suppose $u\bar{g}v^{-1} = \bar{g}$. Then $u'\bar{g}_{11}(v')^{-1} = \bar{g}_{11}$ so that, by induction hypothesis, $u' = v'$ for all but finitely many $\bar{g}_{11} \in M_{n-1,1}$; and

$$\hat{g} = -\bar{g}_{11}v_{12} + u'\bar{g}_{12} + u_{12} = \bar{g}_{12}.$$

Calculating $\hat{g}$ and setting it equal to $\bar{g}_{12}$, we get the equations:

$$\begin{array}{llllll}
1) & -d_{1n} & -a_{12}d_{2n} & +c_{1,n-1}a_{n-1,n} & +c_{1n} & =0 \\
2) & -d_{2n} & -a_{23}d_{3n} & +c_{2,n-1}a_{n-1,n} & +c_{2n} & =0 \\
3) & -d_{3n} & & +c_{3,n-1}a_{n-1,n} & +c_{3n} & =0 \\
4) & -d_{4n} & -a_{45}d_{5n} & +c_{4,n-1}a_{n-1,n} & +c_{4n} & =0 \\
 & \cdot & \cdot & \cdot & \cdot & \cdot \\
n-2) & -d_{n-2,n} & -a_{n-2,n-1}d_{n-1,n} & +c_{n-2,n-1}a_{n-1,n} & +c_{n-2,n} & =0 \\
n-1) & -d_{n-1,n} & & +a_{n-1,n} & +c_{n-1,n} & =a_{n-1,n}.
\end{array}$$

The final equation gives $c_{n-1,n} = d_{n-1,n}$.

Consider the third equation. Since there are no restrictions on $a_{n-1,n}$ (other than it be non-zero) we can choose $a_{n-1,n} \notin \{(d_{3n} - c_{3n})/c_{3,n-1}\}$. Then 3) can be satisfied only if $c_{3,n-1} = 0$, in which case $c_{3n} = d_{3n}$.

Consider, as typical of all remaining equations, equation 1). We can choose a_{12} in such a way that 1) cannot be satisfied unless $d_{2n} = 0 = c_{1,n-1}$, in which case $c_{1n} = d_{1n}$. This is permissible since only *finitely* many choices for a_{12} have been excluded by the induction step.

Proceeding in like manner, it follows that, by keeping $a_{12}, a_{23}, a_{45}, a_{56}, \cdots, a_{n-1,n}$ out of certain finite sets of values, we can assume that $\hat{g} = g_{12}$ implies $u_{12} = v_{12}$. However, for all but finitely many $\bar{g}_{11} \in M_{n-1,1}$, $u'\bar{g}_{11}(v')^{-1} = \bar{g}_{11}$ implies $u' = v'$. Putting the last two statements together, our claim is established. Q. E. D.

We have, therefore, by Lemma 2,

COROLLARY 2. $R(\mathfrak{A}_{nk}) = \mathfrak{A}_{n,k+1}$, $1 \leqq k \leqq n-3$.

We may now state:

THEOREM 1. *For each $n \geqq 4$, $\mathfrak{A}_{n0}$ is a maximal abelian subalgebra of length $n-2$.*

Proof. $\mathfrak{A}_{n0}$ is maximal abelian by Lemma 6.

Since $G_{n0} \subset N_{n1} \subset N_{n2} \subset \cdots \subset N_{n,n-2} = G_n$, statement (3) follows, and, using Lemmas 8 and 9 and Corollary 2, statement (4) follows.

That is, by Definition 3, $\mathfrak{A}_n$ is of length $n-2$. Q. E. D.

THEOREM 2. *For each $n \geqq 4$, $R(\mathfrak{A}_{n0})$ is not a factor.*

Proof. The center, Z_{n1} of N_{n1} consists of all elements of the form

$$\begin{vmatrix} 1 & 0 & a & 0 \cdots 0 & b \\ 0 & 1 & 0 & 0 \cdots 0 & 0 \\ 0 & 0 & 1 & 0 \cdots 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \cdots \cdot & \cdot \\ 0 & 0 & 0 & 0 \cdots 0 & 1 \end{vmatrix}$$

If $\mathfrak{Z}_n$ is the center of $R(\mathfrak{A}_n) = \mathfrak{A}_{n1}$ and $\mathfrak{A}_{Z_{n1}}$ the subalgebra associated with Z_{n1}, then certainly $\mathfrak{A}_{Z_{n1}} \subseteq \mathfrak{Z}_n$ (using Lemma 3 we can, in fact, show that equality holds) so that $\mathfrak{Z}_n$ is non-trivial and $R(\mathfrak{A}_{n0})$ is not a factor. Q. E. D.

THEOREM 3. *For each $n \geqq 4$, $R^2(\mathfrak{A}_{n0})$ is a factor.*

Proof. Since $R^2(\mathfrak{A}_{n0}) = \mathfrak{A}_{n2}$, we must show that each class

$$A_{\bar{g}} = \{g\bar{g}g^{-1} \mid g \in N_{n2}\}$$

is infinite $(\bar{g} \in N_{n2}, \bar{g} \neq e)$.
The proof is the same, mutatis mutandis, as that for Lemma 5.

Combining Theorems 2 and 3, we have

COROLLARY 3. *For each $n \geqq 4$, $\mathfrak{A}_{n0}$ is 2-semi-regular.*

By Lemma 4, the length of a subalgebra is a *-algebraic invariant. Since all hyperfinite factors are *-isomorphic, all $\mathfrak{A}_n$ are *-isomorphic. These two facts, together with Theorem 1 and Corollary 3, serve to establish Theorem I.

4. Theorem II. The proof of Theorem II is exactly analogous to that for Theorem I. The following remarks, however, are called for:

(i) The group G_n remains the same, except that here $n \geqq 5$.

(ii) For each $n \geqq 5$, neither $R(\mathfrak{A}_{n0})$ nor $R^2(\mathfrak{A}_{n0})$ is a factor, but $R^3(\mathfrak{A}_{n0})$ is.

(iii) The abelian subgroup G_{n0} consists of elements of the form

$$\begin{vmatrix} 1 & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \cdots b_{1,n-1} & b_{1n} \\ 0 & 1 & b_{12} & b_{13} & 0 & 0 \cdots 0 & 0 \\ 0 & 0 & 1 & b_{12} & 0 & 0 \cdots 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \cdots 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdots \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 \cdots 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \cdots 0 & 1 \end{vmatrix}$$

(iv) The normalizer, N_{n1}, of G_{n0} consists of those elements $g = |a_{ij}|$ of G_n with $a_{11} = 1$, $a_{23} = a_{12} + x$, $a_{34} = a_{12} + 2x$, $(x \in F)$, $a_{3j} = a_{4j} = 0$ $(j \geqq 5)$.

(v) N_{n2} consists of those elements $g = |a_{ij}|$ of G_n with $a_{11} = 1$, $a_{3j} = 0$ if $5 \leqq j \leqq n-1$, and $a_{4j} = 0$ if $j \geqq 5$.

(vi) N_{nk} $(3 \leqq k \leqq n-2)$ consists of all $g = |a_{ij}|$ of G_n with third and fourth rows as follows:

$$\begin{array}{cccccccccc} 0 & 0 & 1 & b_{34} & 0 & 0 & 0 \cdots 0 & b_{3,n-k+2} & b_{3,n-k+3} \cdots b_{3n} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \cdots 0 & 0 & b_{4,n-k+3} \cdots b_{4n}. \end{array}$$

FORDHAM UNIVERSITY.

REFERENCES.

[1] J. Dixmier, *Les Algèbres d'Opérateurs dans l'Espace Hilbertien*, Paris, Gauthier-Villars, 1957.

[2] ———, "Sous-anneaux abéliens maximaux dans les facteurs de type fini," *Annals of Mathematics*, vol. 59 (1954), pp. 279-286.

[3] F. J. Murray and J. von Neumann, "On rings of operators," *Annals of Mathematics*, vol. 36 (1937), pp. 116-229.

[4] ———, "On rings of operators IV," *Annals of Mathematics*, vol. 44 (1943), pp. 716-808.

[5] L. Pukánszky, "On maximal abelian subrings of factors of type II_1," *Canadian Journal of Mathematics*, vol. 12 (1960), pp. 289-296.

[6] J. Schwartz, *Notes on W^*-algebras*, Lecture Notes, New York, Courant Institute of Mathematical Sciences, 1963.

[7] Sister R. J. Tauer, "Maximal abelian subalgebras in finite factors of type II," *Transactions of the American Mathematical Society*, vol. 114 (1965), pp. 281-308.

[8] J. von Neumann, "On some algebraical properties of operator rings," *Annals of Mathematics*, vol. 44 (1943), pp. 709-715.

STUDIES IN EQUISINGULARITY II.*
EQUISINGULARITY IN CODIMENSION 1 (AND CHARACTERISTIC ZERO).

By Oscar Zariski.

Introduction. We deal with an r-dimensional algebroid hypersurface V which has a singular point at its origin O. We consider the singular locus W of V and we assume that

(A) *W has codimension* 1.

Under this assumption we introduce the concept of equisingularity of V at O, along W. For V to be equisingular at O, along W, we first of all require that

(B) *O be a simple point of W.*

When this condition is satisfied we can speak of *W-transversal sections of V*, at O (Definition 3.3). Any such section is an embedded (see Definition 2.1) algebroid scheme (in the sense of § 1), of dimension 1 or 2 (see Proposition 3.5; the case of dimension 2 may occur, but then only for "special" transversal sections). There is only one W-transversal section V_P of V at the *general* point P of W (see Note (b) at the end of § 3), and V_P is always an embedded algebroid *curve* (hence essentially a plane algebroid curve). We define equisingularity of V at O, along W, by the following condition:

(C) *There exists a W-transversal section V_0 of V, at O, such that V_0 is a curve and such that the two plane algebroid curves V_P and V_0 have equivalent singularities at P and O respectively* (*Definition* 4.1).

Here, equivalence of singularities of plane algebroid curves is intended in the sense defined in our paper [3].

When conditions (A), (B) and (C) are satisfied, we say that V has at its origin O a *singularity of dimensionality type* 1.

Pending further investigation of the concept of equivalence of singularities of plane algebroid curves in the case of characteristic $p \neq 0$ (see [3], Introduction), we restrict ourselves in this paper to the case of ground fields

Received February 5, 1965.

* This research was supported by grants from the National Science Foundation and the Air Force Office of Scientific Research.

which have characteristic zero (and are algebraically closed). It is proved (Corollary 5.3) that if *V has at O a singularity of dimensionality type* 1, *then all W-transversal section of V at O are curves and have equivalent singularities at O.*

The bulk of the paper consists of the derivation of a number of *criteria of equisingularity* in codimension 1 (or—what is the same thing—criteria for singularities of dimensionality type 1). We derive 4 such criteria, namely:

(a) Existence of equisingular local parameters (Definition 4.3 and Theorem 4.4).

(b) A Jacobian criterion (Theorem 5.1).

(c) An inductive criterion, based on the beshavior of V under a monoidal transformation centered at W (Theorem 7.4).

(d) A criterion due to Whitney and Thom (Theorem 8.1).

With these and other results proved in this paper, the theory of equisingularity in codimension 1 can now be regarded as complete (in characteristic zero).

In the complex domain, one can prove (as Whitney does; see [7], § 12; see also [5]) the existence of a "nice" fibration of V, in the neighborhood of W, a fibration induced by a fibration of the ambient affine $(r+1)$-space. In another paper of this series we shall prove a more general theorem, and for this reason we do not discuss, in this paper, the topological implication of equisingularity in the complex domain.

1. Algebroid schemes and varieties. For the purposes of this paper, by an *algebroid scheme* S we mean the spectrum of a complete noetherian equicharacteristic local ring $\mathfrak{o}$. If $k \subset \mathfrak{o}$ is any field of representatives of $\mathfrak{o}$, then we shall say that S is defined over k. The maximal ideal $\mathfrak{m}$ of $\mathfrak{o}$, as a point of the space S, will be called *the origin* of S. Since $\mathfrak{o}$ is always the homomorphic image of a formal power series ring over k, it follows that if S is an algebroid scheme then, for some integer n, we have

$$S = \operatorname{Spec}(k[[X_1, X_2, \cdots, X_n]]/\mathfrak{A}), \tag{1}$$

where $\mathfrak{A}$ is an ideal in the power series ring $k[[X]]$ $(= k[[X_1, X_2, \cdots, X_n]])$ of n indeterminates. We denote by $\mathfrak{o}_S$ the fundamental structure sheaf of S:

$$\mathfrak{o}_{S,x} = \mathfrak{o}_x = \text{ring of quotients of } \mathfrak{o}, \text{ with respect to the prime ideal } x \text{ of } \mathfrak{o},$$

and by $\rho_y{}^x$ the canonical homomorphism of $\mathfrak{o}_x$ into $\mathfrak{o}_y$ $(x, y \in S; x \supset y$, i.e., x is a specialization of y).

The *dimension* of S is defined as the Krull dimension of the local ring $\mathfrak{o}$ (= maximum of the dimensions of the prime ideals of $\mathfrak{A}$). If all the prime ideals of the zero ideal in $\mathfrak{o}$ have the same dimension r, then we say that S is unmixed, of dimension r.

An *algebroid variety* V is a reduced algebroid scheme. If $S = V$ is an alegbroid variety then the zero ideal in $\mathfrak{o}$ is an (irredundant) finite intersection if prime ideals:

$$(0) = \mathfrak{p}_1 \cap \mathfrak{p}_2 \cap \cdots \cap \mathfrak{p}_h. \tag{2}$$

Let V be an algebroid variety. The total ring of quotients K of $\mathfrak{o}$ is then a direct sum of fields:

$$K = \bigoplus_{i=1}^{h} K_i,$$

where h is the integer which occurs in (2), and we have

$$\mathfrak{p}_i = \mathfrak{o} \cap \mathfrak{P}_i,$$

where

$$\mathfrak{P}_i = \bigoplus_{j \neq i} K_j.$$

For any set (α) of distinct indices $\alpha_1, \alpha_2, \cdots, \alpha_q$ $(1 \leqq \alpha_\mu \leqq h)$ we set

$$K_{(\alpha)} = \bigoplus_{\mu=1}^{q} K_{\alpha_\mu} \quad (= \text{set of all elements } a_1 + a_2 + \cdots + a_h \text{ of } K,\ a_i \in K_i, \text{ such that } a_i = 0 \text{ if } i \notin (\alpha)),$$

and we denote by $\phi_{(\alpha)}$ the canonical surjection $K \to K_{(\alpha)}$. If (α) and (β) are two sets of indices and if $(\alpha) \supset (\beta)$, we denote by $\phi_{(\beta)}{}^{(\alpha)}$ the canonical surjection $K_{(\alpha)} \to K_{(\beta)}$. We have

$$\operatorname{Ker} \phi_{(\alpha)} = \bigoplus_{j \notin (\alpha)} K_j = \bigcap_{\mu=1}^{q} \mathfrak{P}_{\alpha_\mu};$$

$$\phi_{(\beta)} = \phi_{(\beta)}{}^{(\alpha)} \circ \phi_{(\alpha)}.$$

If $x \in V$ is a prime ideal of $\mathfrak{o}$, let $\mathfrak{p}_{\alpha_1}, \mathfrak{p}_{\alpha_2}, \cdots, \mathfrak{p}_{\alpha_q}$ be those prime ideals $\mathfrak{p}_i$ in (2) which are contained in x, let $(\alpha) = (\alpha_1, \alpha_2, \cdots, \alpha_q)$ and let

$$\mathfrak{N} = \bigcap_{\mu=1}^{q} \mathfrak{p}_{\alpha_\mu}.$$

Then $\mathfrak{N}$ is the kernel of the canonical homomorphism of $\mathfrak{o}$ into $\mathfrak{o}_x$. Since $\mathfrak{N}$ is also the kernel of $\phi_{(\alpha)} \mid \mathfrak{o}$, we have a canonical injection $\psi_x \colon \mathfrak{o}_x \to K_{(\alpha)}$.

If y is another point of V such that $x \supset y$, and if $(\beta) = (\beta_1, \beta_2, \cdots, \beta_s)$ is the set of indices such that $\mathfrak{p}_i \subset y$ if and only if $i \in (\beta)$ (whence $(\alpha) \supset (\beta)$), then one sees at once that

$$\phi_{(\beta)}{}^{(\alpha)} \circ \psi_x = \psi_y \circ \rho_y{}^x.$$

It follows that we can identify each $\mathfrak{o}_x$ with its ψ_x-image in $K_{(\alpha)}$, and when that is done then $\rho_y{}^x$ becomes identifed with $\phi_{(\beta)}{}^{(\alpha)} \mid \mathfrak{o}_x$. We assume that these identifications have been carried out. So now all the local rings $\mathfrak{o}_x$ $(x \in V)$ are subrings of K, and, for given x, the set (α) of indices, defined above, can be characterized as follows: $K_{(\alpha)}$ is the smallest subring of K which is a sum of direct summands K_i and which contains $\mathfrak{o}_x$.

If $h = 1$, i.e., if K is a field, then the algebroid variety V is *irreducible.* If $h > 1$, each of the h prime ideals $\mathfrak{p}_i$ in (2) defines an irreducible component V_i of V, and we have $V = \bigcup_{i=1}^{h} V_i$. More generally, if $(\alpha) = (\alpha_1, \alpha_2, \cdots, \alpha_q)$ is any set of distinct indices $(1 \leqq \alpha \leqq h)$ then we set $V_{(\alpha)} = \bigcup_{\mu=1}^{q} V_{\alpha_\mu}$. Then it follows at once that $\mathfrak{o}_x \subset K_{(\alpha)}$ if and only if no V_i, $i \notin (\alpha)$, contains x, and that $\mathfrak{o}_x$ can then be identified with $\mathfrak{o}_{V_{(\alpha)},x}$.

2. **Embedded algebroid varieties ("hypersurfaces").** From now on we shall denote the elements (points) of an algebroid scheme S by capital Latin letters. In particular, the origin of S, i.e., the maximal ideal $\mathfrak{m}$ of $\mathfrak{o}$, will be denoted by O. We shall assume from now on that $\mathfrak{o}$ is *not* a regular ring. We denote by r the Krull dimension of $\mathfrak{o}$.

Definition 2.1. *An algebroid scheme $S = \operatorname{Spec}(\mathfrak{o})$, of dimension r, is embedded if the maximal ideal $\mathfrak{m}$ of $\mathfrak{o}$ has a basis of $r+1$ elements.*

If $\mathfrak{m}$ has a basis of $r+1$ elements, any such basis is minimal (since we have assumed that $\mathfrak{o}$ is not a regular ring). Hence S is an embedded scheme if and only if $\dim_k(\mathfrak{m}/\mathfrak{m}^2) = r+1$ (k = any field of representatives of $\mathfrak{o}$). Any minimal basis of $\mathfrak{m}$ will be then called *a system of local coördinates of S at O.*

Let S be embedded, let $x_1, x_2, \cdots, x_{r+1}$ be local coördinates of S at O and let k be a field of representatives of $\mathfrak{o}$, which we fix once and for always. Then $\mathfrak{o}$ is a homomorphic image of a power series ring $k[[X_1, X_2, \cdots, X_{r+1}]]$:

$$\mathfrak{o} = k[[x_1, x_2, \cdots, x_{r+1}]] = k[[X_1, X_2, \cdots, X_{r+1}]]/\mathfrak{B},$$

where $\mathfrak{B}$ is an ideal in $k[[X]]$. Let us now assume that S is unmixed. Since $\dim S = r$, it follows at once that $\mathfrak{B}$ is a *principal ideal* $(f(X_1, X_2, \cdots, X_{r+1}))$,

where $f(X)$ is a non-unit power series in $k[[X_1, X_2, \cdots, X_{r+1}]]$. We shall write symbolically

$$f(X_1, X_2, \cdots, X_{r+1}) = 0, \tag{3}$$

and we shall say that equation (3) represents an embedding of S into the affine $(r+1)$-space A_{r+1}. For a given set of local coördinates $x_1, x_2, \cdots, x_{r+1}$, the power series f is uniquely determined to within an arbitrary unit factor in $k[[X]]$. The transition to another set of local coördinates $y_1, y_2, \cdots, y_{r+1}$ will lead to another embedding of S, obtained from (3) by a formal analytic transformation of affine coördinates, biholomorphic at the origin.

The scheme S defined by (3) is a variety, if and only if f has no multiple factors.

From now on we assume that S is an unmixed embedded variety (denoted by V), of dimension r. We fix once and for always a field k of representatives in $\mathfrak{o}$.

Definition 2.2. *A set of r elements $x_1, x_2, \cdots, x_r$ of $\mathfrak{m}$ is called a system of local parameters of V at O if* (a) *$x_1, x_2, \cdots, x_r$ are parameters of $\mathfrak{o}$* (i.e., *if the ideal generated by $x_1, x_2, \cdots, x_r$ is primary for $\mathfrak{m}$*) *and* (b) *if the set $\{x_1, x_2, \cdots, x_r\}$ can be extended to a system $\{x_1, x_2, \cdots, x_r, x_{r+1}\}$ of local coördinates of V at O.*

If (3) represents an embedding of V in A_{r+1}, relative to a given system of local coördinates $x_1, x_2, \cdots, x_r, x_{r+1}$, then $\{x_1, x_2, \cdots, x_r\}$ is a system of local parameters if and only if $f(0, 0, \cdots, 0, 1) \neq 0$, i.e., if and only if the power series f is regular in X_{r+1}.

Assuming that f is regular in X_{r+1}, we write $f(0, 0, \cdots, 0, X_{r+1}) = cX_{r+1}{}^{\nu}$ + term of degree $> \nu$ $(c \in k, c \neq 0)$. Then ν is the intersection multiplicity of V with the linear space $X_1 = X_2 = \cdots = X_r = 0$; or—in intrinsic terms of local algebra—ν is the multiplicity of the (primary) ideal generated in $\mathfrak{o}$ by the local parameters $x_1, x_2, \cdots, x_r$. We have $\nu > 1$, since $\mathfrak{o}$ is not a regular ring. By the Weierstrass preparation theorem, we can choose the arbitrary unit factor in f in such a way that f becomes a monic polynomial in X_{r+1}, with coefficients in $k[[X_1, X_2, \cdots, X_r]]$:

$$f = X_{r+1}{}^{\nu} + \sum_{i=1}^{\nu} A_i(X_1, X_2, \cdots, X_r) X_{r+1}{}^{\nu-i}, \quad A_i(0, 0, \cdots, 0) = 0. \tag{4}$$

Let s be the degree of the leading form of f:

$$f = f_s(X_1, X_2, \cdots, X_{r+1}) + f_{s+1}(X_1, X_2, \cdots, X_{r+1}) + \cdots,$$

where f_i is a homogeneous polynomial of degree i, with coefficients in k. The integer s is the multiplicity of the local ring $\mathfrak{o}$, or—in geometric terminology—O is an s-fold point of V.

Definition 2.3. A system of local parameters $x_1, x_2, \cdots, x_r$ of V at O is called transversal if the multiplicity of the ideal $\mathfrak{o}x_1 + \mathfrak{o}x_2 + \cdots + \mathfrak{o}x_r$ is equal to the multiplicity s of $\mathfrak{o}$.

Thus, in (4), we have $\nu = s$ if and only if the local parameters $x_1, x_2, \cdots, x_r$ are transversal. In that case (and in that case only) the leading form of each power series $A_i(X_1, X_2, \cdots, X_r)$ $(i = 1, 2, \cdots, \nu)$ is of degree $\geqq i$.

From the general theory of local rings it is known that there always exist transversal local parameters. We shall not elaborate this point, since in the set-up which interests us, the field k will always be infinite (even algebraically closed), and in that case it is obvious that if $\{x_1, x_2, \cdots, x_r, x_{r+1}\}$ is any system of local coördinates then there exists a matrix $\| c_{ij} \|$, with r rows and $r+1$ columns, with elements in k, such that the r elements $x_i' = \sum_{j=1}^{r+1} c_{ij} x_j$ form a system of transversal local parameters.

Let f be of the form (4), with $x_1, x_2, \cdots, x_r$ as local parameters and $x_1, x_2, \cdots, x_r, x_{r+1}$ as local coördinates. We say that $x_1, x_2, \cdots, x_r$ are *separating* local parameters if f is a separable polynomial in X_{r+1} (separating local parameters always exist if k is a perfect field). If $x_1, x_2, \cdots, x_r$ are separating local parameters then the discriminant $D(X_1, X_2, \cdots, X_r)$ of f with respect to X_{r+1} is a non-zero power series and is a non-unit (since $\nu > 1$, and $X_{r+1} = 0$ is a ν-fold root of $f(0, 0, \cdots, 0, X_{r+1})$. This power series depends only (up to a unit factor in $k[[X_1, X_2, \cdots, X_r]]$) on the choice of the local separating parameters $x_1, x_2, \cdots, x_r$ (and not on the choice of x_{r+1}), for if x'_{r+1} is any other element of $\mathfrak{m}$ such that $x_1, x_2, \cdots, x_r, x'_{r+1}$ is a system of local coördinates, then

$$\mathfrak{o} = \sum_{i=0}^{\nu-1} k[[x_1, x_2, \cdots, x_r]] x_{r+1}{}^i = \sum_{i=0}^{\nu-1} k[[x_1, x_2, \cdots, x_r]] x'_{r+1}{}^{\nu},$$

and hence the discriminants of the two bases

$$\{1, x_{r+1}, \cdots, x_{r+1}{}^{\nu-1}\}, \quad \{1, x'_{r+1}, \cdots, x'_{r+1}{}^{\nu-1}\}$$

of $\mathfrak{o}$ over $k[[x_1, x_2, \cdots, x_r]]$ differ only by a unit factor in $k[[x_1, x_2, \cdots, x_r]]$.

We denote by D_0 the product of the *distinct* irreducible factors of D and by Δ the embedded variety in A_r defined by the equation $D_0 = 0$. We call Δ the *critical variety* relative to the local (separating) parameters $x_1, x_2, \cdots, x_r$.

3. Transversal sections. In this section V stands for an unmixed algebroid variety, of dimension r, not necessarily embedded. Let W be an irreducible subvariety of V, let P be the general point of W, and let A be an arbitrary ***simple*** point of W. Let $\rho = \mathrm{cod}_V P$ $(= \mathrm{cod}_V W = \dim \mathfrak{o}_{V,P})$, $\sigma = \mathrm{cod}_V A$ (whence $\sigma \geqq \rho$ and $\dim \mathfrak{o}_{W,A} = \sigma - \rho$). We set

$$\mathfrak{O} = \mathfrak{o}_{V,A},$$
$$\mathfrak{o}' = \mathfrak{o}/P,$$

whence

$$\mathfrak{o}_{W,A} = \mathfrak{O}/\mathfrak{O}P = \mathfrak{o}'_{A/P}. \tag{5}$$

Definition 3.1. *If $y_1, y_2, \cdots, y_{\sigma-\rho}$ are elements of $\mathfrak{m}$ and if $y'_1, y'_2, \cdots, y'_{\sigma-\rho}$ are their P-residues in $\mathfrak{o}'$, then the elements $y_1, y_2, \cdots, y_{\sigma-\rho}$ are called W-transversal parameters of V at A if the elements $y'_1, y'_2, \cdots, y'_{\sigma-\rho}$ form a system of regular parameters of the (regular) local ring $\mathfrak{o}_{W,A}$.*

The existence of W-transversal parameters at A follows from the fact that the element of A/P generate the maximal ideal of $\mathfrak{o}_{W,A}$. It is clear, on the other hand, that any set of W-transversal parameters of V at A consists of elements of A.

In the sequel, if L is any subset of $\mathfrak{o}$ we write $\mathfrak{O}L$ instead of $\mathfrak{O}\cdot\psi(\mathrm{L})$, where ψ is the canonical map of $\mathfrak{o}$ into $\mathfrak{O}$.

COROLLARY 3.2. *If $y_1, y_2, \cdots, y_{\sigma-\rho} \in \mathfrak{m}$, then the y's are W-tranversal parameters of V at A if and only if*

$$\mathfrak{O}\cdot(y_1, y_2, \cdots, y_{\sigma-\rho}) = \mathfrak{O}A. \tag{6}$$

This is an immediate consequence of (5).

Definition 3.3. *Let $y_1, y_2, \cdots, y_{\sigma-\rho}$ be W-transversal parameters of V at A, and let $\mathfrak{O}^*$ be the completion of the local ring $\mathfrak{O}$ $(= \mathfrak{o}_{V,A})$. We set*

$$\mathfrak{O}^*_{(y)} = \mathfrak{O}^*/(\mathfrak{O}^* y_1 + \mathfrak{O}^* y_2 + \cdots + \mathfrak{O}^*_{\sigma-\rho}), \tag{7}$$

$$V^*_{(y)} = \mathrm{Spec}(\mathfrak{O}^*_{(y)}), \tag{7'}$$

and we call the algebroid scheme $V^*_{(y)}$ a W-transversal section of V at the point A (more precisely: the W-transversal section of V at A, relative to the W-transversal parameters $y_1, y_3, \cdots, y_{\sigma-\rho}$).

Note that if we set

$$\mathfrak{O}_{(y)} = \mathfrak{O}/(\mathfrak{O}y_1 + \mathfrak{O}y_2 + \cdots + \mathfrak{O}y_{\sigma-\rho}), \tag{8}$$

$$V_{(y)} = \mathrm{Spec}(\mathfrak{O}_y), \tag{8'}$$

then $\mathfrak{O}^*_{(y)}$ is the completion of the local ring $\mathfrak{O}_{(y)}$, and $V^*_{(y)}$ is the completion of the scheme $V_{(y)}$.

The local ring $\mathfrak{O}^*_{(y)}$ may have nilpotent elements, and thus $V^*_{(y)}$ is not necessarily an algebroid *variety*. All we can say is that $V^*_{(y)}$ is an algebroid scheme of dimension $\geqq \rho$ (since $\dim \mathfrak{O} = \sigma$); it is defined over the $\boldsymbol{k}_A$, where $\boldsymbol{k}_A$ is any field of representatives of the local ring $\mathfrak{O}_{V,A}$. Since k is algebraically closed, there exist fields of representatives $\boldsymbol{k}_A$ which contain the field k, *and it is only such field* $\boldsymbol{k}_A$ *that will be allowed in the sequel.* Later on, in applications, we may want to consider $V^*_{(y)}$ over an algebraic closure $\bar{\boldsymbol{k}}_A$ of $\boldsymbol{k}_A$, by passing from $\mathfrak{O}^*_{(y)}$ to $\mathfrak{O}^*_{(y)} \otimes_{\boldsymbol{k}_A} \bar{\boldsymbol{k}}_A$.

The schemes W and $V_{(y)}$ are subschemes of V, and we can speak of the intersection of their supports. In this sense we have

$$W \cap V_{(y)} = A,$$

as follows at once from Corollary 3.2.

PROPOSITION 3.4. *Assume that* $\mathfrak{o}$ *is a Macaulay ring and that* $y_1, y_2, \cdots, y_{\sigma-\rho}$ *are* W*-transversal parameters of* V *at* A. *If* $V^*_{(y)}$ *is of dimension exactly* ρ, *then* $V^*_{(y)}$ *is unmixed. A necessary and sufficient condition that* $V^*_{(y)}$ *be of dimension* ρ *is that* $\{\psi(y_1), \psi(y_2), \cdots, \psi(y_{\sigma-\rho})\}$ *be a prime sequence in* $\mathfrak{O}$; (*here* ψ *is the canonical homomorphism of* $\mathfrak{o}$ *into* $\mathfrak{O}$), *and a sufficient condition is that* $\{y_1, y_2, \cdots, y_{\sigma-\rho}\}$ *be a prime sequence in* $\mathfrak{o}$.

Proof. Since $\mathfrak{o}$ is a Macaulay ring, also $\mathfrak{O}$ is a Macaulay ring ([4], Theorem 2, Corollary 4, p. 399), and so is $\mathfrak{O}^*$ ([4], Theorem 2, Corollary 6, p. 400). The first two assertions of the proposition are then well-known statements on Macaulay rings. The last statement is a consequence of the easily verified fact that if $\{a_1, a_2, \cdots, a_q\}$ is a prime sequence in $\mathfrak{o}$, consisting of elements of A, then $\{\psi(a_1), \psi(a_2), \cdots, \psi(a_q)\}$ is a prime sequence in $\mathfrak{o}_A$ $(=\mathfrak{O})$.

Example. Let V be the affine cone $X_1X_2 - X_3X_4 = 0$ in $\boldsymbol{A}_4$, let W be the plane $X_1 = X_3 = 0$ (whence $\rho = 1$) and let A be the origin. We have $\mathfrak{o}/P = k[[x_1, x_2, x_3, x_4]]/(x_1, x_3) = k[[x_2, x_4]]$. Hence x_2, x_4 are W-transversal parameters at O. However, the corresponding W-transversal section is the plane $X_2 = X_4 = 0$, hence has dimension > 1 $(= \rho)$. On the other hand, if we set $y_1 = x_2 - x_1$, $y_2 = x_4 - x_3$, then also y_1, y_2 are W-transversal parameters, and this time we get a W-transversal section which is exactly of dimension 1 (namely, the pair of lines $X_1 = X_2 = X_3 = X_4$ and $X_1 = X_2 = -X_3 = -X_4$). Possibly one should restrict the class of W-transversal sections by allowing only such sections which have the right dimension ρ.

PROPOSITION 3.5. *If V is an embedded variety then the maximal ideal of $\mathfrak{O}_{(y)}$ has a basis of $\rho+1$ elements. Hence $V^*_{(y)}$ is either of dimension ρ or $\rho+1$; in the latter case, $\mathfrak{O}^*_{(y)}$ is a regular ring, and in either case $V^*_{(y)}$ is an embedded unmixed scheme (since $\mathfrak{o}$ is a Macaulay ring).*

Proof. If V is embedded, then in the notations of § 2 we have $\mathfrak{o}=k[[X_1,X_2,\cdots,X_{r+1}]](f)$ (and thus $\mathfrak{o}$ is a Macaulay ring). Let $S=\operatorname{Spec}(k[[X_1,X_2,\cdots,X_{r+1}]])$. Then $\mathfrak{o}_{S,A}$ is a *regular* ring of dimension $\sigma+1$. If P' is the prime ideal in $k[[X]]$ which contains f and is mapped onto P, then $\mathfrak{o}_{S,A}/\mathfrak{o}_{S,A}P'$ $(=\mathfrak{o}_{W,A})$ is a *regular* ring of dimension $\sigma-\rho$. It follows at once that $\mathfrak{o}_{S,A}P'$ has a basis of $\rho+1$ elements. Passing to $\mathfrak{o}_{V,A}=\mathfrak{O}=\mathfrak{o}_{S,A}/\mathfrak{o}_{S,A}f$, we conclude that $\mathfrak{O}P$ has a basis of $\rho+1$ elements. It follows therefore from (6) and from the definition (8) of $\mathfrak{O}_{(y)}$, that the maximal ideal of $\mathfrak{O}_{(y)}$ has a basis of $\rho+1$ elements. This completes the proof.

We note the following special cases:

(a) $A=O$ $(=\mathfrak{m})$. In this case $V_{(y)}$ is a scheme defined over k. We have $\sigma=r$, $(y)=(y_1,y_2,\cdots,y_{r-\rho})$, and $\mathfrak{O}=\mathfrak{O}^*=\mathfrak{o}$.

(b) $A=P$ $(=$general point of $W)$. In this case the empty set is the only set of W-transversal parameters (since $\sigma=\rho$), there is only one W-transversal section of V at P, namely $V^*_P=\operatorname{Spec}(\mathfrak{O}^*)$, where $\mathfrak{O}^*$ is the completion of $\mathfrak{o}_{V,P}$; it is an algebroid unmixed *variety*, of dimension ρ, defined over the field $\boldsymbol{k}_P$ $(\cong \mathfrak{o}_{V,P}/\mathfrak{m}_{V,P})$.

4. Definition and a basic criterion of equisingularity in codimension 1 (and characteristic zero). From now on we assume that the ground field k is of characteristic zero and algebraically closed. We assume that V is an embedded algebroid variety of dimension r and that the origin O of V is a singular point of V. We consider an irreducible algebroid subvariety W, of codimension 1 on V, having a simple point at O. There always exist W-transversal sections $V^*_{(y)}$ of V at O, which are of dimension 1 (and not 2), and, by Proposition 3.5, any such a section is an embedded algebroid scheme, hence as we may say, a 1-dimensional "algebroid cycle" in this affine plane, which may have multiple components. $V^*_{(y)}$ is defined over k and has origin O. On the other hand, *the* W-transversal section of V at the general point P of W is an algebroid embedded curve, having origin P and defined over any field of representatives $\boldsymbol{k}_P$ of $\mathfrak{o}^*_{V,P}$; we shall denote this curve of V^*_P. We note that since $k\subset \boldsymbol{k}_P$, $V^*_{(y)}$ is also defined over $\boldsymbol{k}_P$.

Definition 4.1. *V is said to be equisingular at O, along W, if there*

exists a W-transversal section $V^*_{(y)}$ of V at O, such that $V^*_{(y)}$ is a curve and such that $V^*_{(y)}$ and V^*_P have equivalent singularities at O and P respectively. (It is implicit in this definition that O is a simple point of W).

[In regard to this definition, we recall that we have proved in ([3], Section 2, Note) that the equivalence of two plane algebroid curves C, D, defined over some common ground field, is an intrinsic relationship between the local rings of C and D (at their respective origins). Thus, the equivalence (or non-equivalence) of V^*_P and $V^*_{(y)}$ is independent of the choice of the field of representatives $\boldsymbol{k}_P$ of $\mathfrak{o}^*_{V,P}$].

Definition 4.2. *We shall say that V has at the point O a singularity of dimensionality type* 1, *if there exists an irreducible subvariety W of V, of codimension* 1, *such that V is equisingular at O, along W.*

Our aim in this section is to prove a basic criterion of equisingularity in codimension 1. We first give the following definition:

Definition 4.3. *A system of local parameters $x_1, x_2, \cdots, x_r$ of V at O (see Definition* 2.2) *is said to be equisingular, if the critical variety $\Delta_{(x)}$ associated with these parameters (see the end of* § 2) *is a regular algebroid variety (i.e., has a simple point at its origin).*

The criterion is as follows:

THEOREM 4.4. *The following conditions are equivalent:*

(a) *The origin O of V is a singularity of dimensionality type* 1.

(b) *There exist equisingular systems of local parameters of V at O.*

(c) *There exist equisingular systems of transversal local parameters of V at O.*

In the course of the proof of this theorem, several other results will be established, and we summarize these results in the following:

THEOREM 4.5.

(1) *If V is equisingular, at O, along an irreducible subvariety W of codimension* 1, *then*: (1a) *W is the entire singular locus of V (and thus W is uniquely determined), and O is a simple point of W*; (1b) *V is equimultiple at O along W, i.e., if P denotes the general point of W then the multiplicities $m_V(P)$, $m_V(O)$ of V, at P and at O respectively, are equal.*[1]

[1] ***However, equimultiplicity of V at O, along the singular locus W of V*** (always assuming that W is of codimension 1 and has a simple point at O) ***does not imply***

(2) *Let* $\{x_1, x_2, \cdots, x_r\}$ *be a system of local parameters of* V *at* O, *let* π_x *denote the natural (surjective) morphism of* V *into* $\operatorname{Spec}(k[[x_1, x_2, \cdots, x_r]]$ *(determined by the injection of* $k[[x_1, x_2, \cdots, x_r]]$ *into* $\mathfrak{o}$), *and let* Δ_x *be the critical variety relative to the parameters* $x_1, x_2, \cdots, x_r$. *Assume that* $\{x_1, x_2, \cdots, x_r\}$ *is an equisingular system. Then*: (2a) $\pi_x^{-1}\{\Delta_x\}$ *is an irreducible subvariety* W *of* V, *of codimension* 1, V *is equisingular at* O *along* W, *and* $\pi_x \mid W \colon W \to \Delta_x$ *is an isomorphism*; (2b) *if* ξ $(=\xi(x_1, x_2, \cdots, x_r))$ *is a generator of the principal ideal in* $k[[x_1, x_2, \cdots, x_r]]$ *which defines* Δ_x *and if* $\mathfrak{O} = \mathfrak{o}_{V,P}$, *where* P *is the general point of* W, *then the* $\mathfrak{m}_{V,P}$*-primary ideal* (ξ) *in* $\mathfrak{O}$ *and the* $\mathfrak{m}$*-primary ideal* $\mathfrak{o}x_1 + \mathfrak{o}x_2 + \cdots + \mathfrak{o}x_r$ *in* $\mathfrak{o}$ *have the same multiplicity.*

[*Note.* Upon extending $\{x_1, x_2, \cdots, x_r\}$ to a system of local coördinates $x_1, x_2, \cdots, x_r, x_{r+1}$, we realize V as a hypersurface in the affine space A_{r+1}. Let π denote the natural (surjective) morphism (projection) of $\operatorname{Spec}(k[[X_1, X_2, \cdots, X_r, X_{r+1}])$ into $\operatorname{Spec}(k[[X_1, X_2, \cdots, X_r]]$. Then $\pi^{-1}\{\Delta_x\}$ is a (cylindrical) algebroid hypersurface H in A_{r+1}, containing W, and defined by the equation $\xi(X_1, X_2, \cdots, X_r) = 0$, and $\pi^{-1}\{O\}$ is the line $L \colon X_1 = X_2 = \cdots = X_r = 0$. Part (2b) of the theorem can then be stated in terms of intersection multiplicities as follows:

$$i(V \cdot H, W; A_{r+1}) = i(V \cdot L, O; A_{r+1}).] \tag{8}$$

Proof. A) Assume that the origin O of V is a singularity of dimensionality type 1, and let then W be an irreducible subvariety of V, of codimension 1, such that V is equisingular at O along W. Let $(y_1, y_2, \cdots, y_{r-1})$ be a system of W-transversal parameters of V at O such that the corresponding W-transversal section $V^*_{(y)}$ of V is an algebroid *curve* and such that the singularity of $V^*_{(y)}$ at its origin O is equivalent to the singularity of V^*_P at P (see Definition 4.1). We know, by the proof of Proposition 3.5, that $\mathfrak{o}_P$ has a basis of 2 elements. We fix such a basis and we denote its elements by y_r, y_{r+1}. Then by (6), Section 3, we have $\mathfrak{m} = \mathfrak{o} \cdot (y_1, y_2, \cdots, y_{r-1}, y_r, y_{r+1})$, i.e., the $r+1$ elements y_i form a system of local coördinates of V at O. Let

$$f(Y_1, Y_2, \cdots, Y_{r+1}) = 0 \tag{9}$$

be an equation of the corresponding embedding of V in affine $(r+1)$-space. From now we shall identify V with the hypersurface (9).

---

equisingularity of V *at* O, *along* W. For example, the singular locus W of the surface $V \colon z^2 - xy^2 = 0$, is the line $y = z = 0$, and we have $m_V(O) = m_V(W) = 2$. The W-transversal section of V at the general point P of W is a curve with an ordinary double point, but no W-transversal sections of V at O can have an ordinary double point (compare with footnote 4).

The subvariety W is now represented by the linear $(r-1)$-space

$$W: Y_r = Y_{r+1} = 0,$$

and thus $f(Y_1, Y_2, \cdots, Y_{r-1}, 0, 0)$ must be identically zero, since $W \subset V$. As field of representatives $\boldsymbol{k}_P$ of $\mathfrak{o}^*_{V,P}$ we can take the field

$$k_P = k\{\{y_1, y_2, \cdots, y_{r-1}\}\}.$$

The W-transversal section V^*_P of V at P is the plane algebroid curve

$$(10) \qquad \Gamma^y: F^y(Y_r, Y_{r+1}) = 0,$$

where

$$(10') \qquad F^y(Y_r, Y_{r+1}) = f(y_1, y_2, \cdots, y_{r-1}, Y_r, Y_{r+1}).$$

Its origin $\bar{O}$ is the point $Y_r = Y_{r+1} = 0$. The coefficients of $F^{(y)}$ belong to the power series ring $k[[y_1, y_2, \cdots, y_{r-1}]]$ (it is clear that $y_1, y_2, \cdots, y_{r-1}$ are analytically independent over k).

The W-transversal section $V^*_{(y)}$ of V at O is the plane algebroid curve

$$(11) \qquad \Gamma^0: F^0(Y_r, Y_{r+1}) = 0,$$

where

$$(11') \qquad F^0(Y_r, Y_{r+1}) = f(0, 0, \cdots, 0, Y_r, Y_{r+1}).$$

In the terminology of our paper [3; § 6], Γ^0 is a specialization of Γ^y over $y \to 0$ (where $y = (y_1, y_2, \cdots, y_{r-1})$).

Let s be the multiplicity of the curve Γ^y at its origin $\bar{O}$ $(Y_r = Y_{r+1} = 0)$. Since Γ^y and Γ^0 have equivalent singularities at the origin $Y_r = Y_{r+1} = 0$, also Γ^0 must have exactly an s-fold point at the origin. Hence, if $F_s{}^y(Y_r, Y_{r+1})$ is the leading form of F^y, then $F_s{}^0(Y_r, Y_{r+1})$ has to be the leading form of F^0, i.e., we must have $F_s{}^0(Y_r, Y_{r+1}) \neq 0$; in other words: the coefficients of $F_s{}^y(Y_r, Y_{r+1})$ (which are power series in $y_1, y_2, \cdots, y_{r-1}$), do not all vanish at $y = (0)$. This shows *that the leading form f_s of $f(Y_1, Y_2, \cdots, Y_r, Y_{r+1})$ is also of degree s and is independent of $Y_1, Y_2, \cdots, Y_{r-1}$*:

$$(12) \qquad f_s(Y_1, Y_2, \cdots, Y_r, Y_{r+1}) = g_s(Y_r, Y_{r+1}),$$

where g_s is a binary form, of degree s, with coefficients in k.

Thus $s = m_V(O)$. But, by the definition of s, it is clear that $s = m_V(P)$ $= m_{\bar{O}}(\Gamma^y))$, where P is the general point of W. This proves part (1b) of Theorem 4.5 (equimultiplicity of V at O, along W).

B) Upon replacing y_r and y_{r+1} by non-special linear combinations $ay_r + by_{r+1}$, $cy_r + dy_{r+1}$, with coefficients in k, we may assume that the line

$Y_r=0$ is not tangent to Γ^0 at the origin $\bar{O}$. This being assumed, two consequences will follow. In the first place, we will have, by (12),

$$f_s(0,0,\cdots,0,Y_{r+1})=g_s(0,Y_{r+1})=cY_{r+1}{}^s, c\neq 0,$$

and hence $y_1,y_2,\cdots,y_r$ are *transversal local parameters* of V, *at* O. In the second place, since Γ^ν and Γ^0 have equivalent singularities, it follows from Theorem 7, part (b), of [3], that if $D(Y_1,Y_2,\cdots,Y_r)$ is the discriminant of $f(Y_1,Y_2,\cdots,Y_r,Y_{r+1})$ with respect to Y_{r+1}, then $D(y_1,y_2,\cdots,Y_{r-1},Y_r)$ is of the form $\epsilon(y_1,y_2,\cdots,y_{r-1},Y_r)\cdot Y_r{}^N$, where $\epsilon(Y_1,Y_2,\cdots,Y_r)$ is a unit in $k[[Y_1,Y_2,\cdots,Y_r]]$; in other words (see Definition 4.3), $y_1,y_2,\cdots,y_r$ are *equisingular* local parameters. *This proves that* 4.4(a) $\Rightarrow$ 4.4(c).

C) Clearly 4.4(c) implies 4.4(b). Thus, in order to complete the proof of Theorem 4.4 we have only to show that 4.4(b) implies 4.4(a). Now, the implication 4.4(b) $\Rightarrow$ 4.4(a) is certainly included in part 2a of Theorem 4.5. So we shall now prove this part of Theorem 4.5.

Assume then that $x=(x_1,x_2,\cdots,x_r)$ is an equisingular system of local parameters of V at O. Upon extending this system to a system $(x_1,x_2,\cdots,x_r,x_{r+1})$ of local coördinates of V at O we obtain a well defined embedding of V as a hypersurface in A_{r+1}:

$$V: f(X_1,X_2,\cdots,X_r,X_{r+1})=0,$$

where f is a monic polynomial in X_{r+1}, say of degree ν, with coefficients in $k[[X_1,X_2,\cdots,X_r]]$. Furthermore, $X_{r+1}=0$ is a ν-fold root of $f(0,0,\cdots,X_{r+1})$. We denote by R the subring $k[[x_1,x_2,\cdots,x_r]]$ of $\mathfrak{o}$, by K the total ring of quotients of $\mathfrak{o}$, and by $\bar{\mathfrak{o}}$ the integral closure of $\mathfrak{o}$ in K ($\bar{\mathfrak{o}}$ is also the integral closure of R in K).

By assumption, the discriminant D of f, with respect to X_{r+1}, is of the form $\epsilon(X_1,X_2,\cdots,X_r)\cdot[h(X_1,X_2,\cdots,X_r)]^N$, where ϵ is a unit in $k[[X_1,X_2,\cdots,X_r]]$ and h is a power series whose leading form is of degree 1. We can therefore assume, without loss of generality, that h is the power series X_r: thus

$$D=\epsilon(X_1,X_2,\cdots,X_r)X_r{}^N. \tag{13}$$

We now set

$$F^0(X_r,X_{r+1})=f(0,0,\cdots,0,X_r,X_{r+1}).$$

In view of the expression (13) of D, the discriminant of F^0, with respect to X_{r+1}, is not identically zero, whence F^0 has no multiple factors. We set

$$\mathfrak{o}_0=k[[X_r,X_{r+1}]]/(F^0),$$
$$K_0=\text{total ring of quotients of } \mathfrak{o}_0,$$
$$\bar{\mathfrak{o}}_0=\text{integral closure of } \mathfrak{o}_0 \text{ in } K_0.$$

We now apply Theorem 5 of ([3], § 5). By that theorem, we have that $\bar{\mathfrak{o}}$ contains a subring isomorphic with $\bar{\mathfrak{o}}_0$ (which we shall identify with $\bar{\mathfrak{o}}_0$), such that $x_1, x_2, \cdots, x_{r-1}$ are analytically independent over $\bar{\mathfrak{o}}_0$ and such that

$$(14) \qquad \bar{\mathfrak{o}} = \bar{\mathfrak{o}}_0[[x_1, x_2, \cdots, x_{r-1}]].$$

Furthermore, with the above identification, x_r is the F^0-residue of X_r, so that $\mathfrak{o}_0 = k[[x_r]][\zeta]$, where ζ is the F^0-residue of X_{r+1}.

Using these facts, we shall prove first that there exists a (non-unit) power series $\phi(x_1, x_2, \cdots, x_{r-1})$ such that

$$(15) \qquad f(x_1, x_2, \cdots, x_r, 0, X_{r+1}) = [X_{r+1} - \phi(x_1, x_2, \cdots, x_{r-1})]^\nu.$$

We first consider the case in which the algebroid variety V is irreducible. In this case $\bar{\mathfrak{o}}$, and hence also $\bar{\mathfrak{o}}_0$, is an integral domain. Since $\bar{\mathfrak{o}}_0$ is the integral closure of the local domain $k[[x_r, \zeta]]$, of dimension 1, $\bar{\mathfrak{o}}_0$ is the power series ring $k[[t]]$, where $t = \sqrt[\nu]{x_r}$. Thus, by (14), we have now

$$(16) \qquad \bar{\mathfrak{o}} = k[[x_1, x_2, \cdots, x_{r-1}, t]],$$

i. e., $\bar{\mathfrak{o}}$ *is a regular ring of dimension* r [*this shows, incidentally, that the normalization of* V *is a non-singular algebroid variety*; compare with Proposition 4.6 below]. In particular,

$$(17) \qquad x_{r+1} = \sum_{i=0}^{\infty} a_i(x_1, x_2, \cdots, x_{r-1}) t^i, \qquad (t = \sqrt[\nu]{x_r}).$$

The conjugates $x_{r+1}^{(j)}$ of x_{r+1} over $k\{\{x_1, x_2, \cdots, x_{r-1}, x_r\}\}$ are obtained by replacing t in (17) by $\omega^j t$, where ω is a primitive ν-th root of unity. Since $f(x_1, x_2, \cdots, x_r, X_{r+1}) = \prod_{j=0}^{\nu-1} (X_{r+1} - x_{r+1}^{(j)})$, we conclude at once that (15) holds, with $\phi(x_1, x_2, \cdots, x_{r-1}) = a_0(x_1, x_2, \cdots, x_{r-1})$.

Now, if V is not irreducible, we factor f into its irreducible factors:

$$f = f_1 f_2 \cdots f_h.$$

The discriminant of each factor f_μ (with respect to X_{r+1}) is still a power of X_r. Hence, by the irreducible case, we have

$$f_\mu(x_1, x_2, \cdots, x_{r-1}, 0, X_{r+1}) = [X_{r+1} - \phi_\mu(x_1, x_2, \cdots, x_{r-1})]^{\nu_\mu},$$

where f_μ is the degree of ν_μ in X_{r+1} $(\nu = \nu_1 + \nu_2 + \cdots + \nu_h)$, and ϕ_μ is a non-unit power series. Applying (16) to each irreducible f_μ, we see that the splitting field of the polynomial $f(x_1, x_2, \cdots, x_r, X_{r+1})$, over $k\{\{x_1, x_2, \cdots, x_r\}\}$ is a field of the type

$$k\{\{x_1, x_2, \cdots, x_{r-1}, t\}\},$$

where $t^q = x_r$, for some integer q, and that each root of $f(x_1, x_2, \cdots, x_r, X_{r+1})$ belongs to $k[[x_1, x_2, \cdots, x_{r-1}, t]]$. Let, for $\mu \neq \mu'$, $X_{r+1} = \xi_{r+1}$ and $X_{r+1} = \xi'_{r+1}$ be the roots of f_μ and $f_{\mu'}$ respectively. Then

$$\xi_{r+1} = \sum_{i=0}^{\infty} a_i(x_1, x_2, \cdots, x_{r-1}) t^i, \qquad a_0 = \phi_\mu$$

$$\xi'_{r+1} = \sum_{i=0}^{\infty} a'_i(x_1, x_2, \cdots, x_{r-1}) t^i, \qquad a' = \phi_{\mu'}.$$

Since $\xi_{r+1} - \xi'_{r-1}$ must divide the discriminant x_r^N $(= t^{qN})$ in $k[[x_1, x_2, \cdots, x_{r-1}, t]]$, it follows at once that $\phi_\mu = \phi_{\mu'}$. This establishes (15), and proves (in the notations of Theorem 4.5, part 2a) that $\pi_x^{-1}\{\Delta\}$ is the irreducible variety W, of codimension 1, defined by

$$X_r = X_{r+1} - \phi(X_1, X_2, \cdots, X_{r-1}) = 0,$$

and that $\pi_x\colon W \rightarrow \Delta_x$ is an isomorphism.

For simplicity, we replace x_{r+1} by $x_{r+1} - \phi(x_1, x_2, \cdots, x_{r-1})$. Then W is defined by

$$X_r = X_{r+1} = 0.$$

Let $\boldsymbol{k}_x = k\{\{x_1, x_2, \cdots, x_{r-1}\}\}$ and let us consider the two plane algebroid curves

$$\Gamma^x\colon F^x(X_r, X_{r+1}) = 0, \qquad (F^x = f(x_1, x_2, \cdots, x_{r-1}, X_r, X_{r+1}));$$
$$\Gamma^0\colon F^0(X_r, X_{r+1}) = 0, \qquad (F^0 = f(0, 0, \cdots, 0, X_r, X_{r+1})),$$

both defined over $\boldsymbol{k}_x$. It is clear that Γ^x is the W-transversal section of W at the general point $(x_1, x_2, \cdots, x_{r-1}, 0, 0)$ of W, and that Γ^0 is the W-transversal section of V at O, relative to the W-transversal parameter $x_1, x_2, \cdots, x_{r-1}$ of V at O.

In view of the expression (13) of the discriminant D, it follows from Theorem 7, (a) of [3] that Γ^x and Γ_0 have equivalent singularities at the origin $X_r = X_{r+1} = 0$. This shows that V is equisingular at O, along W, and completes the proof of Theorem 4.5, part 2a, and also of Theorem 4.4.

Part 2b of Theorem 4.5 is now obvious, since both multiplicities in question are equal to the degree ν of f in X_{r+1} (the element ξ is now x_r).

COROLLARY 4.6. *Let V have at O a singularity of dimensionality type* 1 *and let W be the singular locus of V (whence W is irreducible, non-singular and of codimension* 1). *Then any W-transversal section of V at O has dimension* 1.

For let $y_1, y_2, \cdots, y_{r-1}$ be W-transversal parameters of V at O. Then we can complete $(y_1, y_2, \cdots, y_{r-1})$ to a system of local coördinates $y_1, y_2, \cdots, y_{r+1}$ of V at O in such a way that the prime ideal of W in $\mathfrak{o}$ is given by $\mathfrak{o}\cdot y_r + \mathfrak{o}\cdot y_{r+1}$. Let $f(Y_1, Y_2, \cdots, Y_{r+1}) = 0$ be the associated embedding of V. Then we know that the leading form f_s of f is a binary form in Y_r, Y_{r+1} (equimultiplicity of V at O, along W). Hence $f(0, 0, \cdots, 0, Y_r, Y_{r+1})$ is not identically zero.

Note. In the course of the above proof we have established the following result:

PROPOSITION 4.7. *If the origin O of V is a singular point of dimensionality type* 1 *and if V has h irreducible components, then the normalization $\bar{V}$ of V consists of h (disjoint) non-singular irreducible algebroid varieties.*[2]

5. A Jacobian criterion of equisingularity. Let $(x_1, x_2, \cdots x_{r+1})$ be a system of local coördinates of V at O, and let

$$f(X_1, X_2, \cdots, X_{r+1}) = 0$$

be the corresponding embedding of V in affine $(r+1)$ space. Consider the following ideal J in $\mathfrak{o}$:

$$J = \mathfrak{o}f'_{x_1} + \mathfrak{o}f'_{x_2} + \cdots + \mathfrak{o}f'_{x_{r+1}}.$$

It is immediately seen that the ideal J is independent of the choice of the local coördinates $x_1, x_2, \cdots, x_{r+1}$; in other words, J is independent of the choice of the embedding of V in A_{r+1}. We call *J the Jacobian ideal of V at O.*

THEOREM 5.1. (*Jacobian criterion of equisingularity*). *Let $\bar{\mathfrak{o}}$ be the integral closure of $\mathfrak{o}$ in the total ring of quotients K of $\mathfrak{o}$. In order that O be a singular point of V of dimensionality type* 1, *it is necessary and sufficient that the following two conditions be satisfied:*

(a) *The ideal $\bar{\mathfrak{o}}J$ is principal.*[3]

(b) *O is a simple point of the singular locus of V.*

In the course of the proof, also the following will be established:

---

[2] *However, V may have a non-singular normalization without necessarily having at O a singularity of dimensionality type* 1. Thus, in the example of footnote 1, the normalization of V is non-singular.

[3] *Condition* (a) *alone does not imply* (b). For instance, consider the surface $z^n = x^a y^b$, with $a \geqq n$, $b \geqq n$. Then $\partial z/\partial x$ and $\partial z/\partial y$ are integral functions of x and y, and hence $\bar{\mathfrak{o}}J = \bar{\mathfrak{o}}f'_z$, where $f = z^n - x^a y^b$. If $n > 1$, then the singular locus of V consists of the two lines $x = z = 0$ and $y = z = 0$.

THEOREM 5.2. *Let $x_1, x_2, \cdots, x_{r+1}$ and f be as above, and assume that V has at O a singularity of dimensionality type* 1 *and that $x_1, x_2, \cdots, x_r$ are local parameters. Then the following conditions are equivalent*:

1) $x_1, x_2, \cdots, x_r$ *are transversal parameters of V at O.*
2) $\bar{o}J = \bar{o}f'_{x_{r+1}}$.

Furthermore, either condition implies that

3) *the local parameters $x_1, x_2, \cdots, x_r$ are equisingular.*[4]

Proof of Theorems 5.1 *and* 5.2. We agree once and for always that if $x_1, x_2, \cdots, x_{r+1}$ are local coördinates of V at O, then $f_{(x_1, x_2, \cdots, x_{r+1})}$ stands for the power series in $k[[X_1, X_2, \cdots, X_{r+1}]]$ such that $f = 0$ represents that embedding of V in A_{r+1} which is determined by the local coördinates x_i (this power series is uniquely determined, to within an arbitrary unit factor).

We first prove the following:

A) *If $\bar{o}J$ is principal, then there exist local coördinates $x_1, x_2, \cdots, x_r, x_{r+1}$ of V at O, such that*:

(18) $\quad x_1, x_2, \cdots, x_r$ *are local transversal parameters of V at O.*

(19) $$\bar{o}J = \bar{o}f'_{x_{r+1}},$$

where $f = f_{(x_1, x_2, \cdots, x_{r+1})}$.

For, start with any system of local coördinates $x_1, x_2, \cdots, x_{r+1}$. If $\| c_{ij} \|$ is any non-singular $(r+1)$-rowed square matrix whose element c_{ij} are in k:

(20) $$| c_{ij} | \neq 0,$$

then the elements $y_1, y_2, \cdots, y_{r+1}$ of $\mathfrak{m}$ defined by

$$x_i = \sum_{j=1}^{r+1} c_{ij} y_j, \qquad i = 1, 2, \cdots, r+1,$$

also constitute a system of local coördinates. We shall impose other inequalities on the c_{ij}.

Let s be the multiplicity of the singular point O of V and let $f_s(X_1, X_2, \cdots, X_{r+1})$ be the leading form of f. We shall require that

(20′) $$f_s(c_{1,r+1}, c_{2,r+1}, \cdots, c_{r+1,r+1}) \neq 0.$$

This inequality insures that $y_1, y_2, \cdots, y_r$ *are local transversal parameters.*

[4] This also shows that the surface V of footnote 1 is not equisingular at O, along the double line $y = z = 0$. For x, y are transversal parameters, without being equisingular parameters (the critical curve $\Delta_{(x,y)}$ is the pair of lines $x = 0$ and $y = 0$.

Let $\bar{\mathfrak{o}}J = \bar{\mathfrak{o}}t$, where $0 \neq t \in \bar{\mathfrak{o}}$, and let

$$\left.\begin{aligned} f'_{x_i} &= \alpha_i t, \quad (i = 1, 2, \cdots, r+1) \\ t &= \sum_{i=1}^{r+1} \beta_i f'_{x_i} \end{aligned}\right\} \alpha_i, \beta_i \in \bar{\mathfrak{o}}.$$

Then

$$(1 - \sum_{i=1}^{r+1} \alpha_i \beta_i) t = 0. \tag{21}$$

We consider in $\bar{\mathfrak{o}}$ the ideal $\mathfrak{A} = (0) : \bar{\mathfrak{o}}t$ and we set $\tilde{\mathfrak{o}} = \bar{\mathfrak{o}}/\mathfrak{A}$. If $\xi \in \bar{\mathfrak{o}}$, we denote by $\tilde{\xi}$ the $\mathfrak{A}$-residue of ξ. We have then by (21):

$$\sum_{i=1}^{r+1} \tilde{\alpha}_i \tilde{\beta}_i = 1. \tag{21'}$$

Let $\tilde{\mathfrak{m}}_1, \tilde{\mathfrak{m}}_2, \cdots, \tilde{\mathfrak{m}}_q$ be the maximal ideals of the semilocal ring $\tilde{\mathfrak{o}}$. By (21′), for any $\nu = 1, 2, \cdots, q$, the α_i are not all in $\tilde{\mathfrak{m}}_\nu$. We impose on the constants $c_{i,r+1}$ the additional conditions

$$\sum_{i=1}^{r+1} c_{i,r+1} \tilde{\alpha}_i \notin \tilde{\mathfrak{m}}_\nu, \qquad \nu = 1, 2, \cdots, q. \tag{20''}$$

With this condition satisfied, the element $\sum c_{i,r+1} \tilde{\alpha}_i$ is a unit in $\tilde{\mathfrak{o}}$. If, then, we set $\epsilon = \sum_{i=1}^{r+1} c_{i,r+1} \alpha_i$, there exists an element ϵ' in $\bar{\mathfrak{o}}$ such that $\epsilon\epsilon' - 1 \in \mathfrak{A}$, i.e., $t = \epsilon\epsilon' t$. Now, let

$$g(Y_1, Y_2, \cdots, Y_{r+1}) = f(\sum_{j=1}^{r+1} c_{1j} Y_j, \sum_{j=1}^{r+1} c_{2j} Y_j, \cdots, \sum_{j=1}^{r+1} c_{r+1,j} Y_j),$$

so that $g(Y_1, Y_2, \cdots, Y_{r+1}) = 0$ is the defining equation of V, relative to the local coördinates $y_1, y_2, \cdots, y_{r+1}$. We have

$$g'_{y_{r+1}} = \sum_{i=1}^{r+1} c_{i,r+1} f'_{x_i} = (\sum_{i=1}^{r+1} c_{i,r+1} \alpha_i) t = \epsilon t.$$

whence $\epsilon' g'_{y_{r+1}} = t$. This shows that $\bar{\mathfrak{o}}J = \bar{\mathfrak{o}} g'_{y_{r+1}}$, and the assertion A) above is proved.

We now prove the following:

B) *Let $x_1, x_2, \cdots, x_{r+1}$ be local coördinates of V at O, satisfying conditions* (18) *and* (19). *Assume furthermore that O is a simple point of the singular locus of V. Then O is a singular point of dimensionality type* 1, *and the transversal local parameters $x_1, x_2, \cdots, x_r$ are equisingular.*

By (19), the ideal $\bar{\mathfrak{o}}J$ is principal. Since we are assuming always that O is not a simple point of V, $\bar{\mathfrak{o}}J$ is not the unit ideal; hence it is unmixed,

of dimension $r-1$. Since every isolated prime ideal of J is the contraction of a prime ideal of $\bar{\mathfrak{o}}J$, it follows the isolated prime ideals of J are of dimension $r-1$. Thus the singular locus of V is of pure dimension $r-1$. Now, in B) we have assumed that this locus is non-singular. It follows that the singular locus of V is an irreducible non-singular variety W, of dimension $r-1$.

We assert that W is the variety of the principal ideal $\mathfrak{o}\cdot f'_{x_{r+1}}$. It is clear that $W \subset \mathfrak{V}(\mathfrak{o}\cdot f'_{x_{r+1}})$. On the other hand, if $\mathfrak{p}$ is any prime ideal of $\mathfrak{o}\cdot f'_{x_{r+1}}$, there exists a prime ideal $\bar{\mathfrak{p}}$ in $\bar{\mathfrak{o}}$ such that $\bar{\mathfrak{p}}\cap\mathfrak{o}=\mathfrak{p}$. For such a prime ideal $\bar{\mathfrak{p}}$ we will have $\bar{\mathfrak{o}}\cdot J=\bar{\mathfrak{o}}\cdot f'_{x_{r+1}}\subset\bar{\mathfrak{p}}$, and hence $J\subset\mathfrak{p}$, showing that $\mathfrak{V}(\mathfrak{p})\subset W$ $(=\mathfrak{V}(J))$, as asserted.

From $W=\mathfrak{V}(\mathfrak{o}\cdot f'_{x_{r+1}})$ follows that $\Delta_x=\pi_x(W)$, where $x=(x_1,x_2,\cdots,x_r)$ and where the notations are the same as in Theorem 4.5, part (2). Since $x_1,x_2,\cdots,x_r$ are transversal parameters at O, the line $X_1=X_2=\cdots=X_r=0$ is not tangent to V, and hence, *a fortiori*, not tangent to W. Since O is a simple point of W, we conclude that the origin $\bar{O}=\pi(O)$ of Δ_x is a simple point of Δ_x. This completes the proof B).

From A) and B), follows the sufficiency of conditions (a) and (b) of Theorem 5.1.

C) Conversely, let us assume that O is a singular point of V, of dimensionality type 1. Let $(x_1,x_2,\cdots,x_r)$ be an equisinguar system of *transversal* local parameters of V at O (see Theorem 4.4, part (c). We shall use the notation of part C) of the proof of Theorems 4.4 and 4.5. We may assume that $X_r=0$ is the critical variety. Then we have (14), and from this it follows that the $r-1$ derivations $\dfrac{\partial}{\partial x_i}$ $(i=1,2,\cdots,r-1)$ are regular on $\bar{\mathfrak{o}}$. In particular, we have

$$\frac{\partial x_{r+1}}{\partial x_i}=-f'_{x_i}/f'_{x_{r+1}}\in\bar{\mathfrak{o}},\qquad i=1,2,\cdots,r-1.$$

If we now prove that also

$$\frac{\partial x_{r+1}}{\partial x_r}\,(=-f'_{x_r}/f'_{x_{r+1}})\in\bar{\mathfrak{o}},$$

then it will follow that $\bar{\mathfrak{o}}J=\bar{\mathfrak{o}}f'_{x_{r+1}}$, and the proof of Theorem 5.1 will be complete.

Let us first consider the case in which V is irreducible. Then we have (17), where $t^\nu=x_r$. Since the parameters $x_1,x_2,\cdots,x_r$ are transversal, we have $\nu=s=$ multiplicity of the singular point O of V. We may also

assume that the singular locus W of V is $X_r = X_{r+1} = 0$; then the term $a_0(x_1, x_2, \cdots, x_{r-1})$ in (17) is missing. Since W is also s-fold for V (Theorem 4.5, part (1b)), we have for every term $cX_1^{i_1}X_2^{i_2}\cdots X_r^{i_r}X_{r+1}^{i_{r+1}}$ $(c \in k)$ in the defining equation of V, the inequality $i_r + i_{r+1} \geqq s$. Since the term X_{r+1}^s occurs in f, it follows at once that in (17) the power series representing x_{r+1} begins with terms of degree $\geqq s$ in t. Since

$$s\frac{\partial x_{r+1}}{\partial x_r}t^{s-1} = \frac{\partial x_{r+1}}{\partial t},$$

it follows that $\partial x_{r+1}/\partial x_r \in \bar{\mathfrak{o}}$, as asserted.

Now, in the general case, if V has h irreducible components $V_1, V_2, \cdots, V_h$, $\bar{\mathfrak{o}}$ is the direct sum of h local domains:

$$\bar{\mathfrak{o}} = \bar{\mathfrak{o}}_1 \oplus \bar{\mathfrak{o}}_2 \oplus \cdots \oplus \bar{\mathfrak{o}}_h,$$

where $\bar{\mathfrak{o}}_j$ is the integral closure of the local domain $\mathfrak{o}_j$ of V_j at O. Let $1 = e_1 + e_2 + \cdots + e_h$ be the decomposition of 1 into mutually orthogonal idempotents ($e_j \in \bar{\mathfrak{o}}$) and let $\phi_j\colon \bar{\mathfrak{o}} \to \bar{\mathfrak{o}}_j$ be the canonical surjection of $\bar{\mathfrak{o}}$ into $\bar{\mathfrak{o}}_j$, defined by $\phi_j(\xi) = e_j\xi$ $(\xi \in \bar{\mathfrak{o}})$. Let $\phi_j(x_i) = x_{ij}$ $(i = 1, 2, \cdots, r+1$; $j = 1, 2, \cdots, h)$. It is clear that for each $j = 1, 2, \cdots, h$, the elements $x_{1j}, x_{2j}, \cdots, x_{rj}$ are *transversal* local parameters of V_j at O, and furthermore these parameters are equisingular [since the critical variety of V_j, relative to $x_{1j}, x_{2j}, \cdots, x_{rj}$, is either empty (if O is a simple point of V_j) or is $X_r = 0$]. Hence, by the irreducible case, we have that

$$\frac{\partial x_{r+1,j}}{\partial x_{i,j}} \in \bar{\mathfrak{o}}_j.$$

Since it is obvious that $\dfrac{\partial x_{r+1}}{\partial x_r} = \sum e_j \dfrac{\partial x_{r+1,j}}{\partial x_{r,j}}$, we conclude that $\dfrac{\partial x_{r+1}}{\partial x_r} \in \bar{\mathfrak{o}}$. This completes the proof of Theorem 5.1.

We note that if $(x_1, x_2, \cdots, x_r)$ is an equisingular system of local parameters of V at O and if these parameters are *not* transversal, then f is a monic polynomial in X_{r+1} of degree ν *greater* than s. If V is irreducible, then the power series in (17) which represents x_{r+1} would begin with a term of degree s in t (assuming, as we may, that the term $a_0(x_1, x_2, \cdots, x_{r-1})$ is missing). Thus $\partial x_{r+1}/\partial x_r$ *would definitely not belong to* $\bar{\mathfrak{o}}$. If V is reducible, then there would have to exist at least one irreducible component V_j of V such that $x_{1j}, x_{2j}, \cdots, x_{rj}$ are *not* transversal local parameters of V_j at O. Then, by the irreducible case we would have

$$\partial x_{r+1,j}/\partial x_{r,j} \notin \bar{\mathfrak{o}}_j,$$

and this would imply that

$$\partial x_{r+1}/\partial x_r \notin \bar{\mathfrak{o}},$$

whence $\bar{\mathfrak{o}}J \neq \bar{\mathfrak{o}}f'_{x_{r+1}}$. We have therefore shown that if $x_1, x_2, \cdots, x_r$ are equisingular local parameters, then $\bar{\mathfrak{o}}J = \bar{\mathfrak{o}}f'_{x_{r+1}}$ if and only if these parameters are transversal. In other words: *if condition* 3) *of Theorem* 5.2 *is satisfied then conditions* 1) *and* 2) *are equivalent.* Assertion B), proved above, says that 1) and 2) together imply 3). So, in order to complete the proof of Theorem 5.2, *we have only to show that conditions* 1) *and* 2) *of that theorem are equivalent.*

We first deal with the case in which V is irreducible. We fix a system $(y_1, y_2, \cdots, y_{r+1})$ of local coördinates of V at O such that the following conditions are satisfied [compare with part C) of the proof of Theorems 4.4 and 4.5]:

a) $y = (y_1, y_2, \cdots, y_r)$ is a system of transversal equisingular parameters.

b) The critical variety Δ_y relative to these parameters is $Y_r = 0$.

c) If

$$g(Y_1, Y_2, \cdots, Y_r, Y_{r+1}) = 0$$

is the embedding of V in an affine $(r+1)$-space, determined by the local coördinates, $y_1, y_2, \cdots, y_r, y_{r+1}$, then the singular locus W of V is defined by $Y_r = Y_{r+1} = 0$; here g is a monic polynomial in Y_{r+1}, of degree $s = m_V(O)$. We have (see (16) and (17)) $\bar{\mathfrak{o}} = k[[y_1, y_2, \cdots, y_{r-1}, t]]$ (where $t^s = y_r$) and

$$(22) \qquad y_{r+1} = a_s(y_1, y_2, \cdots, y_{r-1})t^s + a_{s+1}(y_1, y_2, \cdots, y_{r-1})t^{s+1} + \cdots,$$

since the leading form g_s depends only on Y_r and Y_{r+1}. If we replace y_{r+1} by

$$y_{r+1} - a_s(y_1, y_2, \cdots, y_{r-1})y_r,$$

then we will have $a_s = 0$ in (22), and the leading form g_s of g will be $Y_{r+1}{}^s$. Hence the r partial derivatives

$$\partial y_{r+1}/\partial y_i, \qquad i = 1, 2, \cdots, r,$$

are non-units in $\bar{\mathfrak{o}}$. In other words, the quotients

$$g'_{y_i}/g'_{y_{r+1}}, \qquad i = 1, 2, \cdots, r,$$

are non-units in $\bar{\mathfrak{o}}$. By part C) of the proof we know that $\bar{\mathfrak{o}}J = \bar{\mathfrak{o}}g'_{y_{r+1}}$. We can write

$$y_i = \sum_{j=1}^{r+1} c_{ij}x_j + \text{terms of higher degree}, \qquad (i = 1, 2, \cdots, r+1)$$

where the c_{ij} are in k (and $|c_{ij}|\neq 0$).
The direction $X_1=X_2=\cdots=X_r=0$ is not tangent to V if and only if $c_{r+1,r+1}\neq 0$.

Let

$$f(X)=g(\sum c_{1j}X_j+\cdots,\sum c_{2j}X_j+\cdots,\sum c_{r+1,j}X_j+\cdots).$$

Then

$$f(X)=0$$

is an equation of the embedding of V, relative to the local coördinates $x_1,x_2,\cdots,x_{r+1}$. We have:

$$f'_{x_{r+1}}=\sum_{i=1}^{r}(c_{i,r+1}+\cdots)g'_{y_i}+(c_{r+1,r+1}+\cdots)g'_{y_{r+1}},$$

where the dots stand for non-units in $\mathfrak{o}$. Since $g'_{y_i}/g'_{y_{r+1}}$ is a non-unit in $\bar{\mathfrak{o}}$ for $i=1,2,\cdots,r$, it follows that $\bar{\mathfrak{o}}f'_{x_{r+1}}=\bar{\mathfrak{o}}g'_{y_{r+1}}$ $(=\bar{\mathfrak{o}}J)$, if and only if $c_{r+1,r+1}\neq 0$, i.e., if and only if $x_1,x_2,\cdots,x_r$ are transversal parameters. This completes the proof in the irreducible case.

Now consider the case in which V is reducible. Let $V_1,V_2,\cdots,V_h$ be the irreducible components of V. In the notations of part C) of the proof, the local parameters $x_1,x_2,\cdots,x_r$ are transversal if and only if, for each $j=1,2,\cdots,h$, the local parameters $x_{1j},x_{2j},\cdots,x_{rj}$ of V_j at O are transversal. In view of the equalities

$$\frac{\partial x_{r+1}}{\partial x_i}=\sum_{j=1}^{h}e_j\frac{\partial x_{r+1,j}}{\partial x_{i,j}}, \qquad (i=1,2,\cdots,r)$$

we have $\partial x_{r+1}/\partial x_i\in\bar{\mathfrak{o}}$ if and only if $\partial x_{r+1,j}/\partial x_{i,j}\in\bar{\mathfrak{o}}_j$ for all j. In other words, $\bar{\mathfrak{o}}J=\bar{\mathfrak{o}}f'_{x_{r+1}}$, if and only if

$$\bar{\mathfrak{o}}J_j=\bar{\mathfrak{o}}_jf'_{j;x_{r+1,j}}, \qquad j=1,2,\cdots,h,$$

where J_j is the Jacobian ideal of V_j at O and where f_j is the irreducible factor of f such that $f_j=0$ is the equation of V_j. In view of the irreducible case settled above, this completes the proof of Theorems 5.1 and 5.2.

The following is an important consequence of Theorem 5.2:

COROLLARY 5.3. *If V has at O a singularity of dimensionality type* 1 *and if W is the singular locus of V, then every W-transversal section $V^*_{(y)}$ of V at O is an embedded curve (i.e., an embedded reduced algebroid scheme, of dimension* 1*). Furthermore, all W-transversal sections $V^*_{(y)}$ of V at O have equivalent singularities at O.*

For, let $(y_1,y_2,\cdots,y_{r-1})$ be any system of W-transversal parameters

of V at O. We know already that $V^*_{(y)}$ is embedded (Proposition 3.4) and has dimension 1 (Corollary 4.6). We complete the set $(y_1, y_2, \cdots, y_{r-1})$ to a system of local coördinates $(y_1, y_2, \cdots, y_{r+1})$ in such a manner that $\mathfrak{o}\cdot y_r + \mathfrak{o}\cdot y_{r+1}$ is the prime ideal of W in $\mathfrak{o}$. Let $f(Y_1, Y_2, \cdots, Y_{r+1}) = 0$ be the associated embedding of V in affine $(r+1)$-space. We know then that the leading form f_s of f is a binary form $f_s(Y_r, Y_{r+1})$ in Y_r, Y_{r+1} (this follows from the equimultiplicity of V along W, at O). Without loss of generality we may assume that $f_s(0,1) \neq 0$ (replace y_r, y_{r+1} by "non-special" linear combinations of y_r, y_{r+1}, with coefficients in k). Then $y_1, y_2, \cdots, y_{r-1}, y_r$ are transversal parameters, and therefore—by Theorem 5.2—equisingular. We have here precisely the situation which was reached at the very end of the proof of Theorems 4.4 and 4.5, and this allows us to conclude that $V^*_{(y)}$ is a curve and that this curve has at O a singularity which is equivalent to the singularity which the W-transversal section of V at the general point P of W has at P.

We shall conclude this section by giving a characterization of equisingular local parameters $x_1, x_2, \cdots, x_r$ (under the assumption that it is known already that V has at O a singularity of dimensionality type 1).

Let V have at O a singularity of dimensionality type 1, let W be the singular locus of V and let $x_1, x_2, \cdots, x_r$ be local parameters of V at O. *We shall use the notations of the Note which follows immediately the statement of Theorem* 4.5, except that if $W' = \pi\{W\}$ then we denote by H the cylindrical hypersurface $\pi^{-1}\{W'\}$.

PROPOSITION 5.4. *In order that the local parameters $x_1, x_2, \cdots, x_r$ be equisingular, it is necessary and sufficient that the following equality hold:*

$$i(V\cdot H, W; A_{r+1}) = i(V\cdot L, O; A_{r+1}) \tag{23}$$

Proof. Assume $x_1, x_2, \cdots, x_r$ are equisingular. Since W' is part of the (*non-singular*) critical variety Δ_x and has dimension $r-1$, it follows that $W' = \Delta_x$. So in this case, H has the same meaning in (23) as it does in the cited Note, and (23) is now merely the equality (8).

Conversely, assume (23). Let ν denote the common value of both sides of (23). Then the equation $f(X_1, X_2, \cdots, X_r, X_{r+1}) = 0$ of the hypersurface V is monic in X_{r+1}, of degree ν. Let P be the prime ideal which defines W and let $\mathfrak{o}/P = k[[\xi_1, \xi_2, \cdots, \xi_r, \xi_{r+1}]]$. The fact that $i(V\cdot H, W; A_{r+1}) = \nu$ signifies that ξ_{r+1} is a ν-fold root of the polynomial $f(\xi_1, \xi_2, \cdots, \xi_r; X_{r+1})$. Hence $\xi_{r+1} \in k[[\xi_1, \xi_2, \cdots, \xi_r]]$, say $\xi_{r+1} = a(\xi_1, \xi_2, \cdots, \xi_r)$. Upon replacing x_{r+1} by $x_{r+1} - a(x_1, x_2, \cdots, x_r)$ we may therefore assume that $\xi_{r+1} = 0$. Since

$\dim W' = r-1$, W' is defined by an irreducible equation $g(X_1, X_2, \cdots, X_r) = 0$, where $g \in k[[X_1, X_2, \cdots, X_r]]$, and W is defined by $g=0$, $X_{r+1}=0$. Since O is a simple point of W, g begins with terms of degree 1. So, without loss of generality, we may assume that $g = X_r$ and that W is the linear space $X_r = X_{r+1} = 0$. Upon replacing x_{r+1} by $cx_r + dx_{r+1}$, where c and d are non special constants in k, we may assume that $x_1, x_2, \cdots, x_{r-1}, x_{r+1}$ *are transversal parameters* of V at O (hence equisingular by Theorem 5.2).

To show that $x_1, x_2, \cdots, x_r$ are equisingular parameters we have to show that the associated critical variety Δ_x is the space $X_r = 0$. Now, clearly, it will be sufficient to show that for each irreducible component V_j of V, the critical variety Δ_x^j of V_j, associated with the local parameters $x_{1j}, x_{2j}, \cdots, x_{rj}$ (see part C of the proof of Theorem 5.1), is $X_r = 0$ (or is empty). So *we may assume that V is irreducible* (since the validity of (22) for V implies the validity of the similar equality for each V_j). In that case we have

$$\bar{\mathfrak{o}} = k[[x_1, x_2, \cdots, x_{r-1}t]],$$

where

$$t^s = x_{r+1},$$

and, by (22),

$$x_r = a_\nu(x_1, x_2, \cdots, x_{r-1})t^\nu + a_{\nu+1}(x_1, x_2, \cdots, x_{r-1})t^{\nu+1} + \cdots$$

where the a_i are power series. Since the hyperplane $X_r = 0$ meets V only in W ($X_{r+1} = 0$ a ν-fold proof of $f(X_1, X_2, \cdots, X_{r-1}, 0; X_{r+1})$), it follows that a_ν is a unit in $k[[x_1, x_2, \cdots, x_{r-1}]]$. We have

$$(24) \qquad f'_{x_{r+1}} = -f'_{x_r} \cdot \frac{\partial x_r}{\partial x_{r+1}} = -f'_{x_r} \cdot t^{\nu-s} \text{ times a unit in } \bar{\mathfrak{o}}.$$

Now, by Theorem 5.2, the local parameters $x_1, x_2, \cdots, x_{r-1}, x_{r+1}$ are equisingular. Hence W is the only subvariety of V along which f'_{x_r} is zero (Theorem 4.5, part 2a). Since also t vanishes only on W ($X_{r+1} = 0$ being the critical variety Δ of V, relative to the equisingular parameters $x_1, x_2, \cdots, x_{r-1}, x_{r+1}$), it follows from (24) that $f'_{x_{r+1}}$ vanishes only on W. This shows that Δ_x is just $X_r = 0$ and completes the proof.

6. Generalities on dilatations. Let V be an algebroid variety (in the sense of § 1), and let $\mathfrak{p}_1, \mathfrak{p}_2, \cdots, \mathfrak{p}_h$ be the prime ideals of the zero ideal in $\mathfrak{o}$. Let $\mathfrak{D}$ be an ideal in $\mathfrak{o}$ and let D_0 denote the set of elements of $\mathfrak{D}$ which are not zero divisors (D_0 may be empty; this happens if and only if $\mathfrak{D} \subset \mathfrak{p}_j$, for some $j = 1, 2, \cdots, h$).

Consider the reduced scheme

$$S = \bigcup_{x \in D_0} \operatorname{Spec}(\mathfrak{o}[x^{-1}\mathfrak{D}]).$$

Here, if K denotes the total ring of quotients of $\mathfrak{o}$, $x^{-1}\mathfrak{D}$ stands for the set of all elements y/x of K, where y ranges over $\mathfrak{D}$. The canonical morphism $T: S \rightarrow V$ is called the $\mathfrak{D}$-*dilatation* of V.

If D_0 is empty, then $\mathfrak{o}[x^{-1}\mathfrak{D}]$ is simply $\mathfrak{o}$, $S = V$ and T is the identity.

Assume D_0 is not empty. An elementary argument shows that there exist bases of $\mathfrak{D}$ which are contained in D_0. Let $(u_1, u_2, \cdots, u_q)$ be such a basis. Then it is easily seen that

$$S = \bigcup_{\alpha=1}^{q} \operatorname{Spec}(\mathfrak{o}'_\alpha),$$

where

$$\mathfrak{o}'_\alpha = \mathfrak{o}\left[\frac{u_1}{u_\alpha}, \frac{u_2}{u_\alpha}, \cdots, \frac{u_q}{u_\alpha}\right].$$

If P is any point of V, we denote by $T^{-1}\{P\}$ the set of points P' of S such that $T(P') = P$. It follows immediately from the definition of S that if $P \notin \mathfrak{V}(\mathfrak{D})$ then $T^{-1}\{P\}$ consists of a single point P' and that $\mathfrak{o}_{V,P} = \mathfrak{o}_{S,P'}$ (*biregularity of* T *on* $V - \mathfrak{V}(\mathfrak{D})$).

We denote by S_P the scheme

$$S_P = \bigcup_{P' \in T^{-1}\{P\}} \operatorname{Spec}(\mathfrak{o}_{S,P'}).$$

We set $\mathfrak{O} = \mathfrak{o}_{V,P}$, we denote by ρ the natural homomorphism of $\mathfrak{o}$ into $\mathfrak{O}$ and we set $\rho(u_\alpha) = \bar{u}_\alpha$. Since the kernel of ρ consists entirely of those elements ξ of $\mathfrak{o}$ for which there exists an element η in $\mathfrak{o}$, $\eta \notin P$, such that $\xi\eta = 0$, and since no u_α is a zero divisor, it follows that no $\bar{u}_\alpha$ is a zero divisor in $\mathfrak{O}$. We set

$$\mathfrak{O}'_\alpha = \mathfrak{O}\left[\frac{\bar{u}_1}{\bar{u}_\alpha}, \frac{\bar{u}_2}{\bar{u}_\alpha}, \cdots, \frac{\bar{u}_q}{\bar{u}_\alpha}\right], \qquad (\alpha = 1, 2, \cdots, q) \tag{25}$$

we adjoin a transcendental z_1 to the total ring of quotients of $\mathfrak{O}$, we set $z_\alpha = z_1 \cdot \dfrac{\bar{u}_\alpha}{\bar{u}_1}$ and

$$R' = \mathfrak{O}[z_1, z_2, \cdots, z_q], \tag{25'}$$

whence R' is a homogeneous ring over $\mathfrak{O}$. It is then immediately seen that

$$S_P = \bigcup_{\alpha=1}^{q} \operatorname{Spec}(\mathfrak{O}'_\alpha), \tag{26}$$

or—what is the same thing—

$$S_P = \operatorname{Proj}(R'). \tag{26'}$$

Denote by $\mathfrak{D}_P$ the extension of the ideal $\mathfrak{D}$ to $\mathfrak{O}$, i.e., let $\mathfrak{D}_P = \mathfrak{O}\rho(\mathfrak{D})$. Let $V_P = \operatorname{Spec}(\mathfrak{O})$. Then it follows at once from (26) that S_P *is the transform of* V_P *by the* $\mathfrak{D}_P$-*dilatation*. This dilatation will be denoted by T_P.

From the expression (26′) of S_P it follows that if $\mathfrak{M}$ is the maximal ideal of $\mathfrak{O}$ then the fibre $T^{-1}\{P\}$, as a subspace of S, can be identified with $\operatorname{Proj}(R'/R'\mathfrak{M})$. Thus, *$T^{-1}\{P\}$ is a projective model, defined over the field $\boldsymbol{k}(P)$.*

In particular, S_O is a projective variety defined over k.

It is clear that the closed points of $T^{-1}\{P\}$ (and also of S_P) are those and only those points P' of $T^{-1}\{P\}$ for which $\boldsymbol{k}(P') = \boldsymbol{k}(P)$. In particular, *the closed points O' of S_O are those points O' of S_O for which $\boldsymbol{k}(O') = k$. However, it is easily seen that $S_O = S$* and that consequently *the closed points of S are the points O' such that $\boldsymbol{k}(O') = k$, and that all these points are in $T^{-1}\{O\}$.*

Proof. If $P' \in S$, we may assume that $P' \in \operatorname{Spec}(\mathfrak{o}'_1)$, and that if $\mathfrak{p}'_1$ is the prime ideal of $\mathfrak{o}'_1$ which represents the point P' then

$$\frac{u_i}{u_1} \notin \mathfrak{p}'_1, \text{ for } i = 1, 2, \cdots, n;$$

$$\frac{u_\nu}{u_1} \in \mathfrak{p}'_1, \text{ for } \nu = n+1, n+2, \cdots, q.$$

Then it follows at once that

$$P' \in \operatorname{Spec}(\mathfrak{o}'_i), \text{ for } i = 1, 2, \cdots, n;$$
$$u_j/u_i \text{ is a unit in } \mathfrak{o}_{S,P'}, \text{ for } i, j = 1, 2, \cdots, n;$$
$$u_\nu/u_i \in \mathfrak{m}_{S,P'}, \text{ for } \nu = n+1, n+2, \cdots, q; i = 1, 2, \cdots, n.$$

Let t_i be the $\mathfrak{p}'_1$-residue of u_i/u_1 $(i = 1, 2, \cdots, n)$. If $\mathfrak{p}'_i$ is the prime ideal of $\mathfrak{o}_i'$ which represents the point P' $(i = 1, 2, \cdots, n)$, then it is seen at once that

$$(27) \qquad \mathfrak{o}'_i/\mathfrak{p}'_i = (\mathfrak{o}/\mathfrak{p})\left[\frac{t_1}{t_i}, \frac{t_2}{t_i}, \cdots, \frac{t_n}{t_i}\right], \qquad (i = 1, 2, \cdots, n),$$

where $\mathfrak{p}$ is the prime ideal in $\mathfrak{o}$ which represents the point $P = T(P')$. Since $P' \notin \operatorname{Spec}(\mathfrak{o}'_\nu)$, for $\nu = n+1, n+2, \cdots, q$, the (27) show at once that the closure $\overline{\{P'\}}$ of P' in S' is given by

$$\bigcup_{i=1}^{n} \operatorname{Spec}(\mathfrak{o}/\mathfrak{p}\left[\frac{t_1}{t_i}, \frac{t_2}{t_i}, \cdots, \frac{t_n}{t_i}\right]),$$

and is a projective model over the local domain $\mathfrak{o}/\mathfrak{p}$. This model is reduced to a point if and only if $P = O$ and $\mathfrak{o}'_1/\mathfrak{p}'_1 = \mathfrak{o}/\mathfrak{m} = k$. This completes the proof.

If O' is a closed point of S, the local ring $\mathfrak{o}_{S,O'}$ is noetherian, equicharac-

teristic, and has k as residue field. If $\mathfrak{o}'^*$ denotes the completion of this local ring, then $\mathrm{Spec}(\mathfrak{o}'^*)$ is an algebroid unmixed variety, of dimension r, defined over k. We shall denote this algebroid variety by $S^*_{O'}$ and we shall refer to it as *the completion* of S at O'. It is the completions of S at its various closed points that we will primarily be concerned with in the sequel.

If W is a subvariety ($=$ closed subset) of V, we mean by the *total transform of* W on S (in symbols: $T^{-1}\{W\}$) the set of all $P' \in S'$ such that $T(P') \in W$. It is immediately seen that $T^{-1}\{W\}$ is the underlying space of a closed subscheme of S, namely of

$$\bigcup_{\alpha=1}^{q} \mathrm{Spec}(\mathfrak{o}'_\alpha/\mathfrak{o}'_\alpha\mathfrak{A}),$$

where $\mathfrak{A}$ is any ideal in $\mathfrak{o}$ such that $W = \mathcal{V}(\mathfrak{A})$.

In particular, consider the subvariety $W = F = \mathcal{V}(\mathfrak{D})$. Since $\mathfrak{o}'_\alpha\mathfrak{D} = \mathfrak{o}'_\alpha u_\alpha$, it follows that if P' is any point of $T^{-1}\{F\}$, then $T^{-1}\{F\}$ is defined, locally at P', by the principal ideal $\mathfrak{o}_{S,P'}\cdot u_\alpha$, for a suitable $\alpha = 1, 2, \cdots, q$. Since it is clear that the closed points of $T^{-1}\{F\}$ are also closed in S, it follows that $T^{-1}\{F\}$ *is unmixed, of dimension* $r-1$, *at each of its closed points.*

If W is an *irreducible* sub-variety of V, we define the *proper transform* $T^{-1}[W]$ by

$$T^{-1}[W] = \text{Closure of } T^{-1}\{P\},$$

where P is the general point of W. It is easily seen that $T^{-1}[W]$ is the union of those irreducible components of $T^{-1}\{W\}$ whose general point lies in $T^{-1}\{P\}$.

PROPOSITION 6.1. *Let P be a point of V, let $\mathfrak{O} = \mathfrak{o}_{V,P}$ and let ρ be the natural homomorphism of $\mathfrak{o}$ into $\mathfrak{O}$. If $T^{-1}\{P\}$ is a finite set then there exists a subring $\mathfrak{O}'$ of the total ring of quotients K of $\mathfrak{O}$ such that* (1) $\mathfrak{O} \subset \mathfrak{O}'$ *and $\mathfrak{O}'$ is a finite $\mathfrak{O}$-module*; (2) $S_P = \mathrm{Spec}(\mathfrak{O}')$. [In other words, S_P is dominated by the normalization of $V_P = \mathrm{Spec}(\mathfrak{O})$.]

Proof. This proposition is well-known and is, in fact, a special (and elementary) case of the "main theorem" ([1]). In the general context of schemes, its proof can be found in ([6], II, 6.2). For convenience of the reader, we outline here a proof which is merely an adaptation of an argument found in our paper ([1], pp. 506-508).

We start with an arbitrary basis $(u_1, u_2, \cdots, u_q)$ of $\mathfrak{D}$. We have $S_P = \mathrm{Proj}(R')$, where R' is the homogeneous ring, over $\mathfrak{O}$, defined in (25').

Sinve $T^{-1}\{P'\}$ is the underlying space of $\mathrm{Proj}(R'/R'\mathfrak{M})$ (where $\mathfrak{M}$ is the maximal ideal of $\mathfrak{O}$), the assumption that $T^{-1}\{P\}$ is a finite set is equivalent with the following: *there is only a finite number of prime (homogeneous) ideals in R' which contain $R'\mathfrak{M}$.* We can therefore find in R' a homogeneous element, of some positive degree n, which is not contained in any of the *non-irrelevant prime* (homogeneous) ideals of $R'\mathfrak{M}$. Let η be such an element, and let, say $\eta = f(z_1, z_2, \cdots, z_q)$, where f is a form, of degree n, with coefficients in $\mathfrak{O}$. In view of the above stated property of $R'\mathfrak{M}$ and in view of our choice of η, it follows that the ideal $R'\mathfrak{M} + R'\eta$ is irrelevant (i.e., it contains a power of the ideal $\mathfrak{Z} = R'z_1 + R'z_2 + \cdots + R'z_q$). Now, it is clear that if $\mathfrak{P}$ is any maximal element in the set of all *non-irrelevant* prime homogeneous ideals of R', then $\mathfrak{P} \cap \mathfrak{O} = \mathfrak{M}$, whence $\mathfrak{P} \supset R'\mathfrak{M}$ and thus $\eta \notin \mathfrak{P}$. *It follows that $R'\eta$ is itself an irrelevant ideal.*

Let, then, d be a positive integer such that $\mathfrak{Z}^d \subset R'\eta$. We denote by $\Omega_1(z), \Omega_2(z), \cdots, \Omega_N(z)$ (in some order) the monomials in $z_1, z_2, \cdots, z_q$ of degree d. Then, if $w(z)$ is *any* monomial in $z_1, z_2, \cdots, z_q$ of degree n, we have relations of the form

$$(28) \qquad \Omega_i(z) \cdot w(z) = \eta \sum_{j=1}^{N} a_{ij}\Omega_j(z), \quad (a_{ij} \in \mathfrak{O}) \qquad (i = 1, 2, \cdots, N).$$

Since none of the elements $\bar{u}_1, \bar{u}_2, \cdots, \bar{u}_q$ is a zero division in $\mathfrak{O}$, it follows that also none of the elements $z_1, z_2, \cdots, z_q$ is a zero divisor in R'. Therefore, it follows from (28) that

$$| w(z) - \delta_{ij}a_{ij}\eta | = 0,$$

and hence

$$(29) \qquad \frac{w(z)}{\eta} \; (= \frac{w(\bar{u})}{f(\bar{u})}) \text{ is integral over } \mathfrak{O}.$$

We now observe that if $R'^n = \sum_{j=1}^{\infty} R'_{jn}$, then we also have $S_P = \mathrm{Proj}(R'^n)$. Since $R'^n = \mathfrak{O}[w_1(z), w_2(z), \cdots, w_N(z)]$, where het $w_i(z)$ are the various monomials, of degree n, in $z_1, z_2, \cdots, z_q$, and since $f(z)$ is a linear combination of the $w_i(z)$, with coefficients in $\mathfrak{O}$, we deduce from (29) (which holds for any $w = w_i$) that the following ring $\mathfrak{O}'$ satisfies all the conditions of this proposition:

$$\mathfrak{O}' = \mathfrak{O}\left[\frac{w_1(\bar{u})}{f(\bar{u})}, \frac{w_2(\bar{u})}{f(\bar{u})}, \cdots, \frac{w_{N'}(\bar{u})}{f(\bar{u})}\right]. \qquad \text{Q.E.D.}$$

7. Equisingularity and monoidal dilatations. Let F be an irreducible subvariety of V, and let $\mathfrak{p}$ be the prime ideal of F in $\mathfrak{o}$. By *the monoidal*

transformation of V with center F (or centered at F) we means the $\mathfrak{p}$-dilatation of V.

LEMMA 7.1. *Let F be an irreducible subvariety of V and let $T\colon S \to F$ be the monoidal transformation of V, centered at F. If V is embedded and if O is a simple point of F, then also $S^*_{O'}$ is embedded for any closed point O' of S.*

Proof. If $\rho = \operatorname{cod}_V F$, then it follows from Proposition 3.5 and from the proof of that proposition, that there exists a system of local coördinates $x_1, x_2, \cdots, x_{r+1}$ of V at O such that $(x_1, x_2, \cdots, x_{\rho+1})$ is a basis of $\mathfrak{p}$. (The elements $x_{\rho+2}, x_{\rho+3}, \cdots, x_{r+1}$ are then F-transversal parameters of V at O). Now, if O' is a closed point of S, and if, say,

$$O' \in \operatorname{Spec}(\mathfrak{o}[\frac{x_2}{x_1}, \frac{x_3}{x_1}, \cdots, \frac{x_{\rho+1}}{x_1}]),$$

then $\boldsymbol{k}(O') = k$ and we may assume that $x_i/x_1 \in \mathfrak{m}' = \mathfrak{m}_{S,O'}$, for $i = 2, 3, \cdots, \rho + 1$. It follows immediately that the elements

$$x_1, \frac{x_2}{x_1}, \cdots, \frac{x_{\rho+1}}{x_1}, x_{\rho+2}, \cdots, x_{r+1}$$

form a basis of $\mathfrak{m}'$. This completes the proof.

PROPOSITION 7.2. *With the same assumptions as those of Lemma 7.1, assume furthermore that* $\operatorname{cod} F = 1$. *Then the following conditions are equivalent:*

(1) *V is equimultiple along F, at O.*

(2) *$T^{-1}\{O\}$ is a finite set.*

Proof. Let $x_1, x_2, \cdots, x_r, x_{r+1}$ be local coördinates of V at O, such that (x_r, x_{r+1}) is a basis of the prime ideal of F and such that $x_1, x_2, \cdots, x_r$ are local transversal parameters of V at O. Let

$$f(X_1, X_2, \cdots, X_{r+1}) = 0$$

be the associated embedding of V in an affine $(r+1)$-space; here f is a monic polynomial in X_{r+1}, of degree $s = m_V(O)$. Assume (1), i.e., assume that $m_V(P) = m_V(O) = s$, where P is the general point of F. Then the leading forms of f is a form of degree s, which depends only on X_r, X_{r+1}:

$$f = f_s(X_r, X_{r+1}) + \text{terms of higher degree}.$$

We have then:

$$f = X_{r+1}{}^s + \sum_{i=1}^{s} A_i(X_1, X_2, \cdots, X_r) X_r{}^i X_{r+1}^{s-i};$$

where A_i is a power series in $X_1, X_2, \cdots, X_r$. Clearly x_r is not a zero divisor in $\mathfrak{o}$. Setting $x'_{r+1} = x_{r+1}/x_r$ we find from $f(x_1, x_2, \cdots, x_{r+1}) = 0$:

$$\frac{f(x)}{x_r{}^s} = x'_{r+1}{}^s + \sum_{i=1}^{s} A_i(x_1, x_2, \cdots, x_r) x'^{s-i}_{r+1} = 0.$$

Thus x'_{r+1} is integral over $\mathfrak{o}$. Hence $\mathfrak{o}[\frac{x_{r+1}}{x_r}]$ is a finite $\mathfrak{o}$-module and

$$S = \operatorname{Spec}(\mathfrak{o}[\frac{x_{r+1}}{x_r}]),$$

which proves (2).

Conversely, assume (2). We apply Proposition 6.1. Due to the fact that k is infinite, we can choose as an element η, in the proof of that proposition, an element of the form $az_1 + bz_2$, with a, b in k and $z_1/z_2 = x_r/x_{r+1}$. So we may assume that $\eta = z_1$, and the proof of Proposition 6.1 tells us that $\mathfrak{o}[\frac{x_{r+1}}{x_r}]$ is integral over $\mathfrak{o}$. We may also assume that $x_1, x_2, \cdots, x_r$ are local parameters of V at O, and that consequently $\mathfrak{o}$ is integral over the power series ring $k[[x_1, x_2, \cdots, x_r]]$. We have then a relation of the form

$$\left(\frac{x_{r+1}}{x_r}\right)^n + \sum_{j=1}^{n} B_j(x_1 \cdots x_r) \left(\frac{x_{r+1}}{x_r}\right)^{n-j} = 0,$$

or

$$g(x_1, x_2, \cdots, x_{r+1}) = 0, \text{ where}$$

$$g(X_1, X_2, \cdots, X_{r+1}) = X_{r+1}^n + \sum_{j=1}^{n} B_j(X_1, \cdots, X_r) X_r{}^j X_{r+1}^{s-j}.$$

The leading form of g is of degree n, and depends only on X_r, X_{r+1}. Since f is a factor of g, it follows that also the leading form of f depends only on X_r, X_{r+1}, and this proves (1) and completes the proof of the proposition.

Corollary 7.3. *The assumption being as in Proposition 7.2, let $x_1, x_2, \cdots, x_r, x_{r+1}$ be local coördinates of V at O such that*: (1) (x_r, x_{r+1}) *is the prime ideal of F*; (2) $x_1, x_2, \cdots, x_r$ *are local transversal parameters of V at O. Let $f_s(X_r, X_{r+1})$ be the leading form of f and let, say, m be the number of distinct linear factors of f_s*:

$$f_s(X_r, X_{r+1}) = \prod_{\alpha=1}^{m} (X_{r+1} - c_\alpha X_r)^{\lambda_\alpha}, \quad c_\alpha \in k$$

$$c_\alpha \neq c_\beta \text{ if } \alpha \neq \beta.$$

Let $X'_{r+1} = X_{r+1}/X_r$ *and let*

$$f(X_1, X_2, \cdots, X_r, X_r X'_{r+1}) = X_r^s f'(X_1, X_2, \cdots, X_r, X'_{r+1}).$$

Then f' *factors into* m *distinct factors in* $k[[X_1, X_2, \cdots, X_r]][X'_{r+1}]$:

$$f'(X_1, X_2, \cdots, X_r, X'_{r+1}) = \prod_{\alpha=1}^{m} f'_\alpha(X_1, X_2, \cdots, X_r, X'_{r+1}),$$

where f'_α *is a monic polynomial in* X_{r+1}, *of degree* λ_α, *and* $f'_\alpha(0, 0, \cdots, 0, X'_{r+1}) = (X'_{r+1} - c_\alpha)^{\lambda_\alpha}$. *If we denote by* V'_α *the algebroid variety in the affine space of the variables* $X_1, X_2, \cdots, X_r, X'_{r+1}$ *centered at the point* $O'_\alpha = (0, 0, \cdots, 0, c_\alpha)$ *and defined by* $f'_\alpha(X_1, X_2, \cdots, X_r, X'_{r+1}) = 0$, *then the T-transform* S *of* V *is the union of the* m *(disjoint) algebroid varieties* $V'_1, V'_2, \cdots, V'_m$. *Furthermore,* $T\{O\} = (O'_1, O'_2, \cdots, O'_m)$.

Obvious.

THEOREM 7.4. *Assume that the singular locus* W *of* V *is of codimension* 1 *and has at* O *a simple point. Let* $T\colon S \to V$ *be the monoidal dilatation of* V, *centered at* W, *and let* P *be the general point of* V. *Then* O *is a singularity of dimensionality type* 1 *if and only if the following conditions are satisfied*:

(1) $T^{-1}\{O\}$ *is a finite set, and the number of points in* $T^{-1}\{O\}$ *is the same as that in* $T^{-1}\{P\}$.

(2) *If* $T^{-1}\{O\} = (O'_1, O'_2, \cdots, O'_m)$, *then, in the notations of Corollary* 7.3, *each point* O'_α *is either a simple point of* V'_α, *or is a singularity of* V'_α *of dimensionality type* 1.

Proof. A) Assume that V has at O a singularity of dimensionality type 1. Then V is equimultiple along W, at O (Theorem 4.5, part 1b), and hence $T^{-1}\{O\}$ is a finite set (Proposition 7.2). Let V^*_P be the W-transversal section of V at P. Let $x_1, x_2, \cdots, x_{r-1}, x_r, x_{r+1}$ be as in Corollary 7.3 and let $V^*_{(x)}$ be the W-transversal section of V at O, relative to the W-transversal parameters $x_1, x_2, \cdots, x_{r-1}$. Then $V^*_{(x)}$ and $V^{(*)}_P$ are algebroid *curves* having equivalent singularities at O and P respectively (Corollary 5.3). Furthermore, $V^*_{(x)}$ and V^*_P are defined respectively by the equations

$$f(0, 0, \cdots, 0; X_r, X_{r+1}) = 0,$$
$$f(x_1, x_2, \cdots, x_{r-1}; X_r, X_{r+1}) = 0,$$

both P and O being represented now by the origin in (X_r, X_{r+1})-plane. By Corollary 7.3, the number m of points in $T^{-1}\{O\}$ is the number of distinct

tangent lines of $V^*_{(x)}$. A similar argument shows that the number of points of $T^{-1}\{P\}$ is equal to the number of distinct tangent lines of V^*_P (it is sufficient to observe that—in the notation of § 6—we have $T^{-1}\{P\} = T_P^{-1}\{P\}$ and that T_P is the monoidal transformation of V^*_P, centered at P). Since $V^*_{(x)}$ and V^*_P have equivalent singularities, assertion (1) follows.

To prove assertion (2), we first observe that, by Theorem 5.2, the parameters $x_1, x_2, \cdots, x_r$ are equisingular, and that consequently the discriminant Δ of f with respect to X_{r+1} is a power of X_r (apart from a unit factor). This implies—in the notations of Corollary 7.3—that, for each $\alpha = 1, 2, \cdots, m$, the discriminant Δ'_α of f'_α, with respect to X'_{r+1}, is a power of X_r. Thus $x_1, x_2, \cdots, x_r$ are equisingular parameters of V'_α at O'_α, showing (Theorem 4.4) that O'_α is a singularity of V'_α of dimensionality type 1 (or is a simple point of V'_α).

B) Assume now that conditions (1) and (2) are satisfied. Again we choose local coördinates $x_1, x_2, \cdots, x_{r+1}$ as in Corollary 7.3 and we use the notations of that corollary. Furthermore, we assume that the W-transversal section $V^*_{(x)}$ of V, at O, relative to the W-transversal parameter $x_1, x_2, \cdots, x_{r-1}$, is a *curve*.

The points of $T^{-1}\{P\}$ can be identified with the points $(x_1, x_2, \cdots, x_{r-1}, 0, \xi')$, where ξ' is any of the roots of the polynomial $f'(x_1, x_2, \cdots, x_{r-1}, 0, X'_{r+1})$ (in X'_{r+1}). By condition (1), there exist exactly m such roots. Hence, for each $\alpha = 1, 2, \cdots, m$, the polynomial $f'_\alpha(x_1, x_2, \cdots, x_{r-1}, 0, X'_{r+1})$ has exactly one root $X'_{r+1} = \xi'_\alpha$, necessarily λ_α-fold. So $\xi'_\alpha \in k[[x_1, x_2, \cdots, x_{r-1}]]$. Now, the singular locus of V lies above $X_r = 0$. Hence, the singular locus of V'_α lies above a subvariety of the hyperplane $X_r = 0$, hence is either empty (if O'_α is simple for V'_α) or projects onto $X_r = 0$ (if O'_α is a singularity of dimensionality type one; use condition (2) of the theorem). Since

$$f'_\alpha(0, \cdots, 0, 0, X'_{r+1}) = (X'_{r+1} - c_\alpha)^{\lambda_\alpha},$$

while

$$f'_\alpha(x_1, x_2, \cdots, x_{r-1}, 0, X'_{r+1}) = (X'_{r+1} - \xi'_\alpha)^{\lambda_\alpha},$$

it follows from Proposition 5.4 that $x_1, x_2, \cdots, x_r$ are equisingular parameters of V'_α at O'_α; in other words: the X'_{r+1}-discriminant of f'_α is a power of X_r, apart from a unit factor (this conclusion holds also if O'_α is simple for V'_α). Therefore, also the X_{r+1}-discriminant of f is a power of X_r, showing that $x_1, x_2, \cdots, x_r$ are equisingular parameters of V at O. This completes the proof of the theorem, in view of Theorem 4.4.

We know (Proposition 4.7) that if V has at O a singularity of dimen-

sionality type 1 then the normalization $\bar{V}$ of V is non-singular. We can add to this the following pertinent consequence of Theorem 7.4:

COROLLARY 7.5. *If V has at O a singularity of dimensionality type 1 then the normalization $\bar{V}$ of V can be obtained from V by a sequence of monoidal transformations.*

For, if V'_α is still singular (for some $\alpha = 1, 2, \cdots, m$), its singular locus W'_α is defined by $X_r = 0$, $X'_{r+1} - \phi_\alpha(X_1, X_2, \cdots, X_r) = 0$, where $\phi_\alpha(x_1, x_2, \cdots, x_r) = \xi'_\alpha$. It is easy to see that the W'_α-transveral section $V'^*_{\alpha,(x)}$ of V'_α, at O'_α, relative to the W'_α-transversal parameters $x_1, x_2, \cdots, x_{r-1}$ of V'_α, is the quadratic transform of the W-transversal section $V^*_{(x)}$ of V, at O. We know that all W-transversal section of V at O have equivalent singularities (Corollary 5.3). Thus, with every V which has at O a singularity of dimensionality type 1, there is associated an equivalence class, $C(V)$, of singularities of embedded algebroid curves (which is *not* the class of simple points, as long as O is actually a singular point of V). We have just seen that if $S = \bigcup_{\alpha=1}^{m} V'_\alpha$ is the monodial transform of V, centered at the singular locus W of V, then also V'_α has at its origin O'_α a singularity of dimensionality type 1 (or a simple point), and the set of m equivalence classes $C(V'_1)$, $C(V'_2), \cdots, C(V'_m)$ is the quadratic transform of $C(V)$. From this the corollary follows.

8. The Whitney-Thom criterion. Let V be an algebroid hypersurface in affine A_{r+1}, with origin O. If $V_1, V_2, \cdots, V_h$ are the irreducible components of V, we denote by P_j the general point of V_j and by σ_j the tangent r-space of V_j at P_j. *We assume that the singular locus W of V is of codimension 1 on V and has a simple point at O.* Let L be the tangent $(r-1)$-space of W at O and let τ_j be the r-space determined by L and P_j.

THEOREM 8.1 (*Whitney-Thom*). *The point O is a singularity of V of dimensionality type 1 if and only if the following conditions are satisfied:*

(1) *V is equimultiple along W, at O.*

(2) *In any specialization $(P_j, \sigma_j, \tau_j) \rightarrow (O, \bar{\sigma}, \bar{\tau})$, we have necessarily $\bar{\sigma} = \bar{\tau}$.*

Proof. We choose the local coördinates $X_1, X_2, \cdots, X_{r+1}$ in A_{r+1} at O in such a way that if $x_1, x_2, \cdots, x_{r+1}$ are their traces on V then $x_1, x_2, \cdots, x_r$ are local transversal parameters of V at O, while W is defined by $X_r = X_{r+1}$

$=0$. From the proof of Proposition 7.2 it follows that condition (1) is equivalent to the following:

(1′) $$\frac{x_{r+1}}{x_r} \in \bar{\mathfrak{o}} = \text{integral closure of } \mathfrak{o}.$$

Let $f(X_1, X_2, \cdots, X_{r+1}) = 0$ be the equation of V, and let $f = f_1 f_2 \cdots f_h$, where $f_j = 0$ is the equation of V_j. Let $K = K_1 \oplus K_2 \oplus \cdots \oplus K_h$ be the direct sum decomposition of the total ring of quotients of $\mathfrak{o}$, where $K_j = \boldsymbol{k}(P_j)$, and let ρ_j be the natural surjection of K into K_j. If $x^{(j)}{}_{r+1} = \rho_j(x_{r+1})$, then P_j is the point $(x_1, x_2, \cdots, x_r, x^{(j)}{}_{r+1})$. (Since the restriction of ρ_j to the subfield $\boldsymbol{k}\{\{x_1, x_2, \cdots, x_r\}\}$ of K is an isomorphism, we identify x_i with $\rho_j(x_i)$, for $i = 1, 2, \cdots, r$). The hyperplane σ_j is defined by the equation:

$$\sigma_j \colon \sum_{i=1}^{r} \left(\frac{\partial f}{\partial X_i}\right)_{P_j} (X_i - x_i) + \left(\frac{\partial f}{\partial X_{r+1}}\right)_{P_j} (X_{r+1} - x^{(j)}{}_{r+1}) = 0.$$

On the other hand, the hyperplane τ_j is defined by

$$\tau_j \colon \frac{x^{(j)}{}_{r+1}}{x_r} X_r - X_{r+1} = 0.$$

Upon division by $\dfrac{\partial f}{\partial x^{(j)}{}_{r+1}}$, the equation of σ_j can also be written as follows:

$$\sigma_j \colon \sum_{i=1}^{r} \frac{\partial x^{(j)}{}_{r+1}}{\partial x_i} (X_i - x_i) - (X_{r+1} - x^{(j)}{}_{r+1}) = 0.$$

If we set $\rho_j(\bar{\mathfrak{o}}) = \bar{\mathfrak{o}}_j$, and denote by $\bar{\mathfrak{m}}_j$ the maximal ideal of the local domain $\bar{\mathfrak{o}}_j$, then condition (2) of the theorem is equivalent to the following:

$$\frac{\partial x^{(j)}{}_{r+1}}{\partial x_i} \in \bar{\mathfrak{m}}_j, \quad \text{for } i = 1, 2, \cdots, r-1;$$

$$\frac{\partial x^{(j)}{}_{r+1}}{\partial x_r} - \frac{x^{(j)}{}_{r+1}}{x_r} \in \bar{\mathfrak{m}}_j,$$

for $j = 1, 2, \cdots, h$; or—equivalently, if we denote by $\bar{\mathfrak{m}}$ the intersection of the maximal ideals of the semilocal ring $\bar{\mathfrak{o}}$:

(2′) $$\begin{cases} \dfrac{\partial x_{r+1}}{\partial x_i} \in \bar{\mathfrak{m}}, \quad \text{for } j = 1, 2, \cdots, r-1; \\[2ex] \dfrac{\partial x_{r+1}}{\partial x_r} - \dfrac{x_{r+1}}{x_r} \in \bar{\mathfrak{m}}. \end{cases}$$

Now, (1′) and (2′) together imply at any rate that $\dfrac{\partial x_{r+1}}{\partial x_i} \in \bar{\mathfrak{o}}$ for $i = 1, 2, \cdots, r$,

or—equivalently—that

$$\bar{\mathfrak{o}}J = \bar{\mathfrak{o}}f'_{x_{r+1}},$$

where J is the Jacobian ideal of V at O. By Theorem 5.1, this shows that conditions (1) and (2) imply that V has at O a singularity of dimensionality type 1 (in view of the assumptions made in regard to the singular locus W of V).

To prove the converse, we may assume that V is irreducible, since equisingularity of V at O, along W, implies equisingularity of each V_j at O, along W. In that case, $\bar{\mathfrak{o}}$ is a local domain. We know already that equisingularity of V at O, along W, implies (1). We may assume that $\frac{x_{r+1}}{x_r}$ is a non-unit in $\bar{\mathfrak{o}}$. Then we have only to show that the partial derivatives $\frac{\partial x_{r+1}}{\partial x_i}$ $(i=1,2,\cdots,r)$ are non-units in $\bar{\mathfrak{o}}$. But this is precisely what has been shown in the course of the proof of Theorem 5.2 (where g is to be replaced by f, and $y_1, y_2, \cdots, y_{r+1}$ are to be replaced by $x_1, x_2, \cdots, x_{r+1}$).

REFERENCES.

[1] O. Zariski, "Foundations of a general theory of birational transformations," *Transcations of the American Mathematical Society*, vol. 53 (1943), pp. 496-542.

[2] ———, "Equisingular points on algebraic varieties," *Seminarii dell' Instituto Nazionale di Alta Matematica*, 1962-63 (Rome), pp. 164-177.

[3] ———, "Studies in equisingularity I. Equivalent singularities of plane algebroid curves," *American Journal of Mathematics*, vol. 87 (1965), pp. 507-536.

[4] O. Zariski and P. Samuel, *Commutative Algebra*, vol. II.

[5] O. Zariski, "Seminar on singularities, II." Lecture notes of the Summer Institute on Algebraic Geometry in Woods Hole, Massachusetts (1964), American Mathematical Society.

[6] A. Grothendieck, *Éléments de Géometrie Algébrique*. II.

[7] H. Whitney, "Local properties of analytic varieties," *Differential and Combinatorial Topology*, Princeton University Press (1965).

A REMARK ON MORDELL'S CONJECTURE.

By David Mumford.*

It is somewhat surprising that the systematic evaluation of the heights of rational points on a curve and on its jacobian variety and particularly of their relation to each other should yield any new information. Nonetheless this appears to be the case and the result is described in this article. Although the main theorem is not even a special case of the very fascinating conjecture of Mordell, still it is an estimate that already reveals that rational points on curves of genus at least 2 are much harder to come by than on curves of genus 0 or 1. It is a quantitative limitation on the heights of such points which is well-known to be false in the case of genus 0 or 1. Incidentally, there is a good explanation why an estimate of this type can be obtained so cheaply, whereas Mordell's conjecture itself could not: namely, results obtained by our methods will more or less automatically apply to the analogous "function field" case [where the ground field is a function field in one variable over a finite field, rather than an algebraic number field]. And in this case, unless further restrictions are imposed, there are curves of any genus with an infinite number of rational points whose heights increase exactly at the rate which we will find.

Let k be an algebraic number field of finite degree over $\boldsymbol{Q}$. Let C be a non-singular projective curve over k of genus g at least 2. Mordell's conjecture asserts that the set of k-rational points on C is finite. Now suppose that a projective embedding of C is fixed, allowing us to talk of the heights, $ht(x)$, of k-rational points of x. Then my result is this:

Theorem. *There are real constants a and b, $a > 0$, such that if the countable set of k-rational points of C is ordered by increasing height—call the points $x_1, x_2, \cdots$—then*

$$ht(x_i) \geqq e^{ai+b}.$$

Because of the well-known properties of heights, this result is not affected by changing the projective embedding of C. An example of the theorem is given by Fermat's curve:

Received February 25, 1965.

* This research was partially supported by the AMS 1964 Summer Institute in Algebraic Geometry and NSF-GP3512.

COROLLARY. *Let $(\alpha_i, \beta_i, \gamma_i)$ be an infinite set of distinct positive integral solutions of the equation*

$$X^n + Y^n = Z^n$$

such that α_i, β_i, γ_i have no common factors and such that $\{\gamma_i\}$ is an increasing sequence. Assume $n \geqq 4$. There are real constants a and b, $a > 0$, such that

$$\gamma_i \geqq e^{(e^{ai+b})}.$$

A final word: that the proof of the theorem appears in as natural and simple a form as it does is due to the collaboration of John Tate; that it appears in print, needless to say, is not.

1. The theory of heights. We fix an algebraic number field k, of finite degree over $\boldsymbol{Q}$. The main result of the "classical" Theory of Weil (cf. [1] and [4]) is the construction of a set of functions as follows:
Given: a scheme X, projective over k, and an element $\delta \in Pic(X)$.
Constuct: a real-valued function on the set of k-rational points X_k, written

$$h_\delta(x),\ x \in V_k$$

In fact, h_δ is not constructed precisely, but only the class of all functions, differing from one member of this class by a *bounded* function is constructed. This construction has the following properties (where $O(x)$ denotes a bounded function of x):

a) If $f\colon X \to Y$ is a k-morphism of schemes X and Y as above, and if $\delta \in Pic(Y)$, then

$$h_\delta(f(x)) = h_{f^*(\delta)}(x) + O(x)$$

b) If $\delta_1, \delta_2 \in Pic(X)$, for X as above, then

$$h_{\delta_1+\delta_2}(x) = h_{\delta_1}(x) + h_{\delta_2}(x) + O(x)$$

c) If D is an effective Cartier divisor on the projective scheme X, and if D defines the element $\delta \in Pic(X)$, then there is a real constant K such that

$$h_\delta(x) \geqq K, \text{ all } x \in X - \text{Support }(D).$$

d) If $\delta \in Pic(X)$ is ample, then for all constants K, the set of points $x \in X_k$ such that $h_\delta(x) \leqq K$ is finite.

The lack of a really definite height function is one of the most awkward aspects of this theory. In case X is assumed to be an abelian variety, this

defest has been remedied by Néron and Tate (cf. [2], [3], [4½]). The simplest way to state their result is this:

THEOREM. *Let X be an abelian variety, and let $\delta \in Pic(X)$. Then the class of functions h_δ on X contains a "quadratic" function on X, i.e., a function f satisfying the identity:*

$$f(x+y+z)-f(x+y)-f(x+z)-f(y+z)$$
$$+f(x)+f(y)+f(z)-f(0)=0.$$

One checks immediately that a real-valued bounded quadratic function is constant. Therefore, if we put the two requirements on h_δ that (1) it is quadratic, and (2) it is 0 at the identity point e, then we obtain a completely well-defined height function. Moreover, we get the important Corollary:

COROLLARY. 1) *If X is an abelian variety, and $\delta_1, \delta_2 \in Pic(X)$, then the normalized height functions on X satisfy:*

$$h_{\delta_1+\delta_2}(x)=h_{\delta_1}(x)+h_{\delta_2}(x), \text{ all } x \in X_k.$$

2) *If $f\colon X \to Y$ is any morphism of abelian varieties, and $\delta \in Pic(Y)$, then*

$$h_{f^*\delta}(x)=h_\delta(f(x))-h_\delta(f(e)),$$

all $x \in X_k$. In particular, if f is a homomorphism (i.e., takes the identity to the identity), then

$$h_{f^*\delta}(x)=h_\delta(f(x)).$$

2. The set-up derived from a curve. We shall assume given a non-singular projective curve C, over k, with genus $g \geqq 1$. The purpose of this section is to give a thorough account of the auxiliary varieties associated to C, the canonical divisor classes that they carry, and their universal properties. For the sake of simplicity, we also assume that a base point $x_0 \in C_k$ has been chosen once and for all; and that all other schemes X occurring in the discussion have base points p_X. (The base points on abelian varieties will be assumed to be their identity points). A general concept which is central to the discussion is the following:

Definition. Let X and Y be connected algebraic schemes over k. A *divisorial correspondence* on $X \times Y$ is an element $\delta \in Pic(X \times Y)$ which is 0 restricted to either of the subschemes $X \times \{p_Y\}$ of $\{p_X\} \times Y$.

First of all, let J be the connected component of the identity of the

Picard scheme of C: i. e., the so-called "Jacobian variety" of C. It is an abelian variety of dimension g. Moreover, J is characterized by the existence of a canonical divisorial correspondence

$$\delta_1 \in Pic(C \times J)$$

which has the universal mapping property (cf. [5] and [6]):

$$(*) \qquad \left\{\begin{array}{l}\text{For all connected algebraic schemes } X, \text{ and all} \\ \text{divisorial correspondences } \eta \text{ on } C \times X, \text{ there is} \\ \text{a unique morphism } f\colon X \to J \text{ such that} \\ \qquad\qquad (1_C \times f)^*(\delta_1) = \eta.\end{array}\right.$$

Secondly, on the non-singular surface $C \times C$ the Weil divisor

$$\Delta - C \times \{x_0\} - \{x_0\} \times C$$

defines an element $\Delta^* \in Pic(C \times C)$ which is clearly a divisorial correspondence. By the $UMP(*)$, there is a unique morphism

$$\phi\colon C \to J$$

such that $\Delta^* = (1_C \times \phi)^*(\delta_1)$.

Thirdly, let $\hat{J}$ be the connected component of the identity of the Picard scheme of J: i. e., the dual abelian variety. $\hat{J}$ is characterized by the existence of a canonical divisorial correspondence

$$\delta_2 \in Pic(J \times \hat{J})$$

which has the universal mapping property:

$$(**) \qquad \left\{\begin{array}{l}\text{For all connected algebraic schemes } X, \text{ and all} \\ \text{divisorial correspondences } \eta \text{ on } J \times X, \text{ there is} \\ \text{a unique morphism } f\colon X \to \hat{J} \text{ such that} \\ \qquad\qquad (1_J \times f)^*(\delta_2) = \eta.\end{array}\right.$$

Fourthly, the morphism ϕ dualizes to a morphism $\hat{\phi}\colon \hat{J} \to J$. Namely, apply the Universal mapping property $(*)$ with $X = \hat{J}$, $\eta = (\phi \times 1_{\hat{J}})^*(\delta_2)$. This means that we get a diagram:

$$(***) \qquad \begin{array}{ccc} C \times \hat{J} & \xrightarrow{\phi \times 1_{\hat{J}}} & J \times \hat{J} \\ \Big\downarrow{\scriptstyle 1_C \times \hat{\phi}} & & \\ C \times J & & \end{array}$$

such that δ_1 and δ_2 induce the same correspondence on $C \times \hat{J}$.

Fifthly, recall the general construction by which divisor classes η on abelian varieties X define homomorphisms from X to its dual $\hat{X}$. There are three maps from $X \times X$ to X—the group law μ and the two projections p_1 and p_2. Then one checks that for any $\eta \in Pic(X)$, the divisor class

$$\mu^*(\eta) - p_1^*(\eta) - p_2^*(\eta)$$

is a divisorial correspondence on $X \times X$. Therefore, by definition of $\hat{X}$, there is a unique morphism $f: X \to \hat{X}$ such that

$$\mu^*\eta - p_1^*\eta - p_2^*\eta = (1_X \times f)^* \begin{bmatrix} \text{canonical class} \\ \text{on } X \times \hat{X} \end{bmatrix}.$$

We will denote f by $\Lambda(\eta)$. Recall that Λ is itself a homomorphism: $\Lambda(\eta_1 \pm \eta_2) = \Lambda(\eta_1) \pm \Lambda(\eta_2)$. In terms of this definition, the central result concerning jacobians is the following (due to Weil [7]).

THEOREM. *$\exists$ an ample divisor Θ on J such that*

$$\hat{\phi} = -\Lambda(\Theta)^{-1}.$$

In fact, recall that Θ is nothing but the sum of the subset $\phi(C)$ in J with itself (with respect to the group law in J) $(g-1)$ times. For reference we write the meaning of this Theorem out as follows:

$$\begin{cases} \psi = -\hat{\phi}^{-1} \\ \text{class of } \underbrace{\mu^*\Theta - p_1^*\Theta - p_2^*\Theta}_{\text{call this } \theta} = (1_J \times \psi)^*(\delta_2). \end{cases}$$

The net result of all this is the following: suppose we identify J with $\hat{J}$ via the isomorphism ψ, or $\Lambda(\Theta)$. Then we have defined the canonical divisor classes:

$$\begin{aligned} &\text{On } C \times C: \Delta^* \\ &\text{On } C \times J: \delta_1 \\ &\text{On } J \times J: \theta = \text{class of } \mu^*\Theta - p_1^*\Theta - p_2^*\Theta \\ &\qquad\qquad\; = \delta_2 \\ &\qquad\qquad\;\; \text{via our} \\ &\qquad\quad\;\; \text{identifications} \end{aligned}$$

These are related by the equations

(a) $\Delta^* = (1_C \times \phi)^*(\delta_1)$

(b) $\delta_1 = -(\phi \times 1_J)^*(\theta)$.

hence

(c) $\Delta^* = -(\phi \times \phi)^*(\theta)$.

Proof. (a) has been pointed out before, and (c) follows from (a) and (b). As for (b), first use the fact that $\Lambda(-\Theta) = -\Lambda(\Theta) = -\psi$. Therefore

$$-\theta = (1_J \times (-\psi))^*\delta_2 = (1_J \times \hat{\phi}^{-1})^*\delta_2.$$

Hence

$$-(1_J \times \hat{\phi})^*\theta = \delta_2$$

and finally:

$$\begin{aligned}(1_C \times \hat{\phi})^*\delta_1 &= (\phi \times 1_{\hat{J}})^*\delta_2 \qquad \text{(This is (***))}\\ &= -(\phi \times 1_{\hat{J}})^*(1_J \times \hat{\phi})^*\theta\\ &= -(\phi \times \hat{\phi})^*\theta\\ &= -(1_C \times \hat{\phi})^*(\phi \times 1_{\hat{J}})^*\theta.\end{aligned}$$

Since $1_C \times \hat{\phi}$ is an isomorphism, (b) follows. Q. E. D.

3. The basic estimates. Once again, we consider a curve C over a number field k, as above. Now we will use the maps obtained in § 2 to obtain properties of the height functions introduced in § 1. The most important height function is $h_\theta(x,y)$ defined for $x, y \in J_k$.

PROPOSITION 1. *$h_\theta(x,y)$ is a symmetric, bilinear form on $J_k \times J_k$. Moreover it is positive definite on J_k/mod torsion.*

Proof. Let $f_1: J \to J \times J$ be the homomorphism mapping x to $x \times e$, and let f_2 map x to $e \times x$. Since θ is a divisorial correspondence, $f_1^*\theta = f_2^*\theta = 0$. Therefore

$$h_\theta(x,e) = h_\theta(f_1(x)) = h_{f_1^*\theta}(x) = 0,$$
$$h_\theta(e,x) = h_\theta(f_2(x)) = h_{f_2^*\theta}(x) = 0.$$

But this means that h_θ is a quadratic function on the product of two groups which is 0 on both factors alone. It is easy to check that this implies that h_θ is bilinear.

Let $\xi: J \times J \to J \times J$ be the morphism mapping $x \times y$ to $y \times x$. Then clearly $\xi^*\theta = \theta$, hence

$$h_\theta(x,y) = h_\theta(\xi(y,x)) = h_{\xi^*\theta}(y,x) = h_\theta(y,x).$$

To evaluate $h_\theta(x,x)$, let $\Delta: J \to J \times J$ be the diagonal morphism, and let $\lambda_2: J \to J$ be multiplication by 2. Then

$$\begin{aligned} h_\theta(x,x) &= h_\theta(\Delta(x)) \\ &= h_{\Delta^*\theta}(x) \\ &= h_{\Delta^*(\mu^*\Theta - p_1^*\Theta - p_2^*\Theta)}(x) \\ &= h_{\lambda_2^*\Theta}(x) - 2h_\Theta(x). \end{aligned}$$

since $\lambda_2 = \mu \circ \Delta$, $1_J = p_i \circ \Delta$. On the other hand, if D is any divisor on J, let D' be the divisor obtained by reflecting D in the origin. Then $\lambda_2^*(D)$ is in the same divisor class as $3D + D'$. Therefore,

$$\begin{aligned} h_\theta(x,x) &= h_\Theta(x) + h_{\Theta'}(x) \\ &= h_\Theta(x) + h_\Theta(-x) \end{aligned}$$

I claim that if this is not positive, then x must be a torsion point on J. Namely, assume $h_\theta(x,x) \leqq 0$. Then for all integers n,

$$\begin{aligned} h_\Theta(nx) + h_\Theta(-nx) &= h_\theta(nx, nx) \\ &= n^2 h_\theta(x,x) \\ &\leqq 0, \end{aligned}$$

hence either $h_\Theta(nx) \leqq 0$ or $h_\Theta(-nx) \leqq 0$. This means that if x is not a torsion point, there are an infinite number of distinct points x_i such that $h_\Theta(x_i) \leqq 0$. Since Θ is ample, this contradicts property (d) of heights. Q. E. D.

By the Mordell-Weil theorem, J_k is a finitely generated abelian group. In particular

$$X = J_k \otimes \boldsymbol{R}$$

is a finite-dimensional real vector space. Moreover, h_θ makes it into a Euclidean space: we will abbreviate the norm $h_\theta(x,y)$ to $\langle x, y\rangle$. The inner product $\langle x, y\rangle$ can be used to compute other heights too:

PROPOSITION 2. *Let $\eta \in Pic(C)$ be a divisor class of degree 0. Then there is a unique point $\tilde{\eta} \in J_k$ such that η equals the restriction of δ_1 to $C \times \{\tilde{\eta}\}$, and*

$$\langle \phi x, \tilde{\eta}\rangle = -h_\eta(x) + O(x), \text{ all } x \in C_k.$$

Proof. The first assertion is part of the definition of the jacobian J of C. The second is an immediate consequence of (b), § 2:

$$\begin{aligned} \langle \phi x, \tilde{\eta}\rangle &= h_\theta(\phi x, \tilde{\eta}) \\ &= h_{(\phi\times 1_J)^*\theta}(x, \tilde{\eta}) + O(x) \\ &= -h_{\delta_1}(x, \tilde{\eta}) + O(x) \\ &= -h_\eta(x) + O(x). \end{aligned}$$

Q. E. D.

PROPOSITION 3. $\langle \phi x, \phi y \rangle = -h_{\Delta^*}(x,y) + O(x,y)$.

Proof. This follows from (c), § 2. Q.E.D.

COROLLARY 1. *There is a constant K such that for $x, y \in C_k$, $x \neq y$,*

$$\langle \phi x, \phi y \rangle \leqq h_{x_0}(x) + h_{x_0}(y) + K.$$

Proof. Recall that $\Delta^* = \Delta - (x_0) \times C - C \times (x_0)$. Apply property (c), § 1 of heights to $h_\Delta(x,y)$; note that the divisor $(x_0) \times C$ (resp. $C \times (x_0)$) is of the form $p_1^*(x\ )$ (resp. $p_2^*(x_0)$); hence $h_{(x_0)\times C}(x,y)$ equals $h_{x_0}(x)$ to within a bounded function and $h_{C\times(x_0)}(x,y)$ equals $h_{x_0}(y)$ to within a bounded function. Q.E.D.

COROLLARY 2. *There is a divisor class $\kappa \in Pic(C)$ of degree 0 such that for $x \in C_k$,*

$$\langle \phi x, \phi x \rangle = 2g h_{x_0}(x) + h_\kappa(x) + O(x).$$

Proof. The self-intersection number (Δ^2) of the diagonal on $C \times C$ is well-known to be $2 - 2g$. Therefore the divisor class on Δ obtained by restricting the class of Δ^* has degree $-2g$. Let

$$f \colon C \to C \times C$$

be the diagonal map. Then there is a divisor class $\kappa \in Pic(C)$ of degree 0 such that

$$f^*(\Delta^*) = -(2gx_0 + \kappa).$$

Therefore

$$\begin{aligned} \langle \phi x, \phi x \rangle &= -h_{\Delta^*}(f(x)) + O(x) \\ &= -h_{f^*(\Delta^*)}(x) + O(x) \\ &= 2g h_{x_0}(x) + h_\kappa(x) + O(x). \end{aligned} \qquad \text{Q.E.D.}$$

Putting Proposition 2 and Corollary 1 and 2 together, we obtain the basic estimate:

There is a constant K, and an element $\tilde{\kappa} \in J_k$ such that if $x, y \in C_k$, $x \neq y$, then

$$\langle \phi x, \phi y \rangle \leqq 1/2g\{\langle \phi x, \phi x \rangle + \langle \phi x, \tilde{\kappa} \rangle + \langle \phi y, \phi y \rangle + \langle \phi y, \tilde{\kappa} \rangle\} + K.$$

4. A packing argument. From here on, we have only to make some elementary observations about Euclidean geometry. First of all, define a new map:

$$C_k \xrightarrow{\psi} X$$

via $\psi(x) = \phi(x) + \dfrac{\bar{\kappa}}{2g-2}$. One checks easily that ψ has the property:

$$\begin{cases} \text{There is a constant } K_2 \text{ such that if } x, y \in C_k,\ x \neq y, \text{ then} \\ \langle \psi x, \psi y \rangle \leqq 1/g \left[\dfrac{\langle \psi x, \psi x \rangle + \langle \psi y, \psi y \rangle}{2} \right] + K_2. \end{cases}$$

Let $\| z \| = \sqrt{\langle z, z \rangle}$, let $f(s) = 1/2(s + 1/s)$, and let

$$\cos(u, v) = \langle u, v \rangle / \| u \| \cdot \| v \|$$

be the cosine of the angle between points $u, v \in X$ in the given norm. Then we can rewrite the above formula as:

$$\cos(\psi x, \psi y) \leqq \frac{1}{g} f \left(\frac{\| \psi x \|}{\| \psi y \|} \right) + \frac{K_2}{\| \psi x \| \cdot \| \psi y \|}$$

Now arrange the countable set of points C_k in a sequence so that

$$\| \psi x_1 \| \leqq \| \psi x_2 \| \leqq \cdot \cdot \cdot .$$

Note that as $\| \psi x \| \sim \sqrt{2gh_{x_0}(x)}$, (Cor. 2, § 3), it follows that $\| \psi x_i \| \to +\infty$ as $i \to \infty$. The following "packing" lemma is well-known:

LEMMA. *There is an integer N such that if $A_1, \cdots, A_N$ are any non-zero elements of X, then for some pair of integers $1 \leqq i, j \leqq N$,*

$$\cos(A_i, A_j) \geqq \tfrac{2}{3}.$$

COROLLARY. *If $g \geqq 2$ and $\| \psi x_n \| > \sqrt{12 K_2}$, then $\| \psi x_{n+N} \| \geqq \frac{5}{3} \| \psi x_n \|$.*

Proof. If not, whenever $n \leqq i \leqq j \leqq n + N$ then $1 \leqq \| \psi x_j \| / \| \psi x_i \| < \frac{5}{3}$. Hence

$$1 \leqq f \left(\frac{\| \psi x_j \|}{\| \psi x_i \|} \right) < 7/6,$$

and

$$\cos(\psi x_j, \psi x_i) < 7/6g + \frac{K_2}{\| \psi x_i \| \cdot \| \psi x_j \|} < \tfrac{2}{3}.$$

This contradicts the lemma. Q. E. D.

COROLLARY. *If $g \geqq 2$, then there are real constants a and b, $a > 0$, such that*

$$\| \psi x_n \| \geqq e^{an+b}.$$

It is now easy to argue backwards and show that $\| \phi x_n \|$, and $ht_{x_0}(x_n)$, and finally $ht_\delta(x_n)$—for any $\delta \in Pic(C)$ of positive degree—also increase exponentially. This will be left to the reader.

REFERENCES.

I. On the theory of heights:

[1] A. Weil, "Arithmetic on algebraic varieties," *Annals of Mathematics*, vol. 53 (1951), p. 412.

[2] S. Lang, "Les formes bilinéaires de Néron et Tate," *Séminaire Bourbaki*, Expose 274 (1964).

[3] A. Néron, "Hauteurs sur les variétés abéliennes" (to appear).

[4] S. Lang, *Diophantine Geometry*, Interscience-Wiley, N. Y., 1962.

[4½] J. Manin, "The Tate height of points on an abelian variety, its variants and applications," *Izvestia Academii Nauk*, vol. 28 (1964), p. 1363.

II. On the theory of Picard schemes and abelian varieties:

[5] S. Lang, *Abelian varieties*, Interscience-Wiley, N. Y., 1959.

[6] A. Grothendieck, *Fondéments de la géométrie algébrique*, Collected Bourbaki talks, Paris, 1962.

[7] A. Weil, *Variétés Abéliennes et Courbes Algébriques*, Hermann & Cie, Paris, 1948.

FINDING A BOUNDARY FOR AN OPEN MANIFOLD.

By W. Browder, J. Levine and G. R. Livesay.*

Given an open manifold, when is it the interior of a compact manifold with boundary? We consider the case of piecewise linear (combinatorial) or smooth manifolds W of dimension $\geqq 6$ and we give necessary and sufficient conditions on the homology of W and homotopy at ∞ (see § 1) so that W is isomorphic (i. e., combinatorially equivalent or diffeomorphic) to the interior of a compact manifold with 1-connected boundary (Theorem 1). As a consequence we get an h-cobordism theorem for certain types of open manifolds of dimension $\geqq 6$ (Corollary to Theorem 2). We are indebted to E. H. Connell for suggesting the latter theorem and the possibility of deducing it from the first.

All manifolds will be piecewise linear or smooth and isomorphic will mean either combinatorially equivalent or diffeomorphic.

1. Statement of results. A space X is said to be 1-connected at ∞ if for any compact $C \subset X$ there is a compact D, $C \subset D \subset X$ such that $X - D$ is 1-connected.

THEOREM 1. *Let W be an open manifold of dimension $\geqq 6$. Then W is isomorphic to the interior of a compact manifold U with 1-connected boundary if and only if the homology $H_*(W)$ is finitely generated and W is 1-connected at ∞. Further such a U is unique up to isomorphism.*

Actually the proof given here could be modified slightly so that the condition of 1-connected at ∞ could be weakened to W having a finite number of ends, each of which is 1-connected (see [8]).

Theorem 1 can be considered as a partial generalization of the result of Stallings [9] that contractible open manifolds of dimension $\geqq 5$ which are 1-connected at ∞ are isomorphic to R^n (the interior of the n-ball).

Two connected manifolds M_1, M_2 (not necessarily closed) are called h-cobordant if there is a manifold with boundary V, with $\partial V = M_1 \cup (-M_2)$, and such that each component M_i of ∂V is a deformation retract of V.

Received December 2, 1964.

* Browder was partially supported by an NSF grant, and also received support from the DSIR at the Topology Symposium in Cambridge, England. Levine was an NSF Postdoctoral Fellow at the Topology Symposium. Livesay was partially supported by an NSF grant.

THEOREM 2. *Let M_1, M_2 satisfy the hypothesis of Theorem* 1, *and let V be an h-cobordism between them which is* 1-*connected at* ∞. *If $\bar{M}_1$, $\bar{M}_2$ are the compact bounded manifolds produced by Theorem* 1, *then $\bar{M}_1$ and $\bar{M}_2$ are h-cobordant, i.e., there is a compact manifold with boundary $\bar{V}$ such that $\partial\bar{V} = \bar{M}_1 \cup \bar{U} \cup \bar{M}_2$, where $\bar{U}$ is an h-cobordism between $\partial\bar{M}_1$ and $\partial\bar{M}_2$, and such that the inclusions $\bar{M}_i \subset \bar{V}$ are homotopy equivalence, $i = 1, 2$.*

COROLLARY. *Let M_1, M_2, V be as above, and in addition $\pi_1 M_1 = \pi_1 M_2 = 0$. Then M_1 is isomorphic with M_2.*

2. Uniqueness of the boundary.

THEOREM 3. *Let U_1 and U_2 be compact n-manifolds with* 1-*connected boundaries, U_1 embedded in interior U_2 and the inclusion of U_1 in U_2 induces homology isomorphism. Suppose also that V is* 1-*connected, where $V = U_2 -$* interior *U_1. Then V is an h-cobordism between ∂U_1 and ∂U_2.*

Proof. The map $(V, \partial U_1) \subset (U_2, U_1)$ is an excision so that $H_*(V, \partial U_1) \cong H_*(U_2, U_1) = 0$ since the inclusion induces homology isomorphism between U_1 and U_2. Since V, ∂U_1, ∂U_2 are 1-connected it follows that ∂U_1 and V are homotopy equivalent by the theorem of J. H. C. Whitehead and it follows that V and ∂U_2 are homotopy equivalent similarly, using relative Poincaré duality.

COROLLARY. *If $W =$ interior $U_1 =$ interior U_2, ∂U_1, ∂U_2* 1-*connected, then ∂U_1 is h-cobordant to ∂U_2.*

Actually the corollary holds in more generality without the assumption of 1-connectedness.

3. Proof of Theorem 1.

PROPOSITION 4. *Let W be as in Theorem* 1. *Then given a compact set C there is a connected manifold U with boundary $U \subset W$, ∂U* 1-*connected, $C \subset$ interior U, such that the inclusion induces a homology isomorphism.*

Before we prove this, we will indicate how Theorem 1 follows from the proposition.

Proof of Theorem 1. Let $C_1 \subset C_2 \subset \cdots \subset W$ be a sequence of compact sets such that $W = \bigcup_{i=1}^{\infty} C_i$, and $W - C_i$ is 1-connected. By Proposition 4 we may find compact manifolds with boundary U_i such that $U_i \supset U_{i-1} \cup C_i$, ∂U_i 1-connected and $U_i \subset W$ induces homology isomorphism. Then $\cup U_i \supset \cup C_i = W$, so $\cup U_i = W$. Set $V_i = \overline{U_{i+1} - U_i}$. Now $V_i \subset W - C_i$, which is 1-connected, $\partial V_i = \partial U_{i+1} \cup \partial U_i$, which are 1-connected, so that it follows from

van Kampen's theorem that $\pi_1(W - C_i) =$ free product of $\pi_1(U_i - C_i), \pi_1(V_i)$ and $\pi_1(W - U_{i+1})$. Since $\pi_1(W - C_i)$ is trivial, it follows that $\pi_1(V_i)$ is trivial. By Theorem 3, V_i is an h-cobordism between ∂U_i and ∂U_{i+1}, and since ∂U_i is 1-connected and has dimension $\geqq 5$, it follows from the h-cobordism theorem of Smale [7] (which has been proved in the piecewise linear case by Stallings) that $\overline{U_{i+1} - U_i}$ is isomorphic to $\partial U_i \times I$. Hence, it follows that each U_i is isomorphic to U_1 and $W = \cup\, U_i$ is isomorphic to interior of U_1, which completes the proof of Theorem 1.

Now we outline the proof of Proposition 4.

Since $H_*(W)$ is finitely generated we may find a compact set $C' \subset W$ such that $H_*(C')$ maps onto $H_*(W)$. Hence for any subset O such that $C' \subset O \subset W$, $H_*(O)$ maps onto $H_*(W)$. We shall show (Lemma 6) that we can find a compact manifold with boundary $U \subset W$ such that (1) $C \cup C' \subset$ interior U, (2) ∂U is 1-connected, and (3) $W - U$ is 1-connected. Set $V =$ closure $W - U$.

Consider the commutative diagram with exact rows:

$$\begin{array}{ccccccccc}
\to H_{k+1}(W) & \xrightarrow{j} & H_{k+1}(W,U) & \xrightarrow{\partial} & H_k(U) & \xrightarrow{i} & H_k(W) \to \\
\uparrow g_2 & & \uparrow g_3 & & \uparrow g_1 & & \uparrow g_2 \\
\to H_{k+1}(V) & \xrightarrow{j'} & H_{k+1}(V,\partial U) & \xrightarrow{\partial'} & H_k(\partial U) & \xrightarrow{i'} & H_k(V) \to .
\end{array}$$

Now $(V, \partial U) \subset (W, U)$ is an excision so that g_3 is an isomorphism. Hence i onto implies i' onto since $j \equiv 0$ implies $j' \equiv 0$. Similarly i' mono implies i mono, since $\partial' \equiv 0$ implies $\partial \equiv 0$. Since i is onto by construction, i' is onto, and it suffices to make i' mono in order to make i an isomorphism.

Now the manifold U obtained from Lemma 5 may in fact have too much homology above dimension 1, so that ∂U will also, and $\ker i'$ will be non-zero. Therefore we shall show how to enlarge U to a larger U' such that the kernel from $H_*(U')$ to $H_*(W)$ is smaller. One way in which to do this is to add handles to U along ∂U to kill some of the excess homology of U, i.e. find $D^k \times D^{n-k} \subset V$, $D^k \times D^{n-k} \cap U = S^{k-1} \times D^{n-k} \subset \partial U$, $S^{k-1} \times 0$ representing an element $x \in H_{k-1}(\partial U)$ which goes to O in $H_{k-1}(V)$, and take $U' = U \cup D^k \times D^{n-k}$.

Assuming $H_i(\partial U) \to H_i(V)$ is mono (hence isomorphism) for $i < k-1$, and onto for $i = k-1$, and since ∂U and V are 1-connected, the relative Hurewicz Theorem implies that $\pi_k(V, \partial U) = H_k(V, \partial U)$ and we get a commutative diagram

$$\begin{array}{ccccccccc}
0 \to \pi_k(V, \partial U) & \longrightarrow & \pi_{k-1}(\partial U) & \longrightarrow & \pi_{k-1}(V) \to 0 \\
\downarrow \cong & & \downarrow & & \downarrow \\
0 \to H_k(V, \partial U) & \longrightarrow & H_{k-1}(\partial U) & \xrightarrow{i'} & H_{k-1}(V) \to 0
\end{array}$$

so that $\ker i'$ is represented by spherical homology classes in the lowest dimension. If $k-1 < \frac{1}{2}(n-1)$, $n = \dim W$, then one may deduce from a general position argument that an element $x \in \ker i' \subset H_{k-1}(\partial U)$ is represented by an embedded $S^{k-1} \subset \partial U$ and using either general position or Whitney's embedding theorem that it bounds an embedded $D^k \subset V$, from which we may get a handle. In case $\frac{1}{2}(n-1) \leqq k-1 < n-3$, we may obtain the handles differently, using Smale's theory of handle bodies (see [7]). At the very last stage, $k-1 = n-3$, still another technique must be used, which is not obviously the same as attaching k dimensional handles.

We give the details in the next sections.

4. Proof of Proposition 4; codimensions > 3. In this section, we will prove a weaker form of Proposition 4.

PROPOSITION 5. *Let W^n be an open manifold of dimension $n \geqq 6$, 1-connected at ∞ and with $H_*(W)$ finitely generated. Then given a compact $C \subset W$ and $k \leqq n-3$, there is a compact manifold with boundary U, with ∂U and $W - U$ 1-connected, $C \subset$ interior of U, and if $i\colon U \to W$ is the inclusion, then $i_*\colon H_i(U) \to H_i(W)$ is an isomorphism for $i < k$, and onto for all i.*

Proposition 4 is the same but without restriction on k.

LEMMA 6. *If W is a manifold of dimension $\geqq 5$ which is 1-connected at ∞ and given compact $C \subset W$, then there exists a compact manifold U with 1-connected boundary with $C \subset$ interior of U and such that $W - U$ is 1-connected.*

(Note that this shows that if W is 1-connected at ∞ of dimension $\geqq 5$, then $\pi_1(W)$ is finitely generated, since by Van Kampen's theorem $\pi_1(W) = \pi_1(U)$.)

Proof of Lemma 6. Let D be compact, $C \subset D \subset W$ such that $W - D$ is 1-connected. We may find a compact manifold with boundary U' with $D \subset$ interior U'. This follows in the differentiable case by choosing a proper function f, with $f(D) = 0$, $f \geqq 0$ and letting $U' = f^{-1}([0, \epsilon])$ where ϵ is a regular value of f. In the piecewise linear case, D lies in a finite subcomplex K of W and we take U' to be a regular neighborhood of K in W. By taking connected sums along the boundary of the different components of U', we

may assume U' is connected. Then $\partial U'$ divides W into two parts, U' and $W - U'$, and U' is connected. Since W is 0-connected at ∞ it follows that all but one of the components of $W - U'$ are compact. Define U'' to be the union of U' and all the compact components of $W - U'$. Since W is connected each such component meets U', so that U'' is connected and $W - U''$ is connected. Then the components of $\partial U''$ may be joined by disjoint arcs in $W - U''$. We let $U''' \equiv U'' \cup$ (closed neighborhood of these arcs) and it follows that U''' and $\partial U'''$ are connected.

Now Lemma 6 follows from:

LEMMA 7. *Let M^n be a closed submanifold of W^{n+1}, $n \geqq 4$, and suppose that $\pi_1(W) = 0$ and $\pi_1(M)$ is finitely generated. Then we can do surgery on M inside W to get M' with $\pi_1(M') = 0$. In particular we can find 2-disks, $D_1^2, \cdots, D_k^2 \subset W$ with $D_i^2 \cap M = S_i^1 =$ boundary of D_i^2, meeting M transversally, such that $M \cup D_1 \cup \cdots \cup D_k$ is simply connected, so that the surgeries corresponding to $S_1^1, \cdots, S_k^1$ produce a simply connected manifold.*

For a proof see [1, Lemma 3.1].

As indicated in § 3, since $H_*(W)$ is finitely generated we may find a compact $C' \subset W$ such that $H_*(C')$ maps onto $H_*(W)$, so that for $C' \subset O \subset W$, $H_*(O)$ maps onto $H_*(W)$. Then applying Lemma 6, we may find $U_1 \subset W$ with ∂U_1 1-connected and $C \cup C' \subset$ interior of U_1, and $W - U_1$ 1-connected.

Also as indicated in § 3, we may kill the kernel of $H_i(\partial U_1) \to H_i(V_1)$ where $V_1 =$ closure of $W - U_1$, and this will kill the kernel of $H_i(U_1) \to H_i(W)$. Since ∂U_1 and V_1 are 1-connected the Hurewicz theorem tells us that every element $x \in H_2(\partial U_1)$ is represented by a map $f\colon S^2 \to \partial U_1$ and if $i_{1*}x = 0$ in $H_2(V_1)$, $i_1\colon \partial U_1 \to V_1$ then f is homotopic to a constant in V_1. Since dimension $\partial U_1 = n - 1 \geqq 5$, it follows from general position that f is homotopic to an embedding g of $S^2 \subset \partial U_1$, and if $n > 6$, g extends to an embedding $\bar{g}\colon D^3 \subset V_1$, $\bar{g} \mid S^2 = g$, and $\bar{g}D^3$ meets ∂U_1 transversally in $g(S^2)$. If $n = 6$, the existence of $\bar{g}\colon D^3 \to V_1$ with the required properties follows from Whitney's embedding theorem [11] or from the result of Irwin [4]. Now for a generator of kernel i_{1*} take such a 3-disk D^3, and define $U_1' =$ regular neighborhood of $U_1 \cup D^3$ in W. This can be made smooth using a theorem of Hirsch [3], or one can define $U_1' = U_1 \cup D^3 \times D^{n-3}$ and round the corners in an appropriate way, where $D^3 \times D^{n-3}$ is a neighborhood of D^3 in V_1. Now $W - U_1'$ is homotopy equivalent to $V_1 - D^3$, and since D^3 is of codimension $\geqq 3$, it follows that $W - U_1'$ is still 1-connected. Similarly $V_1 \cap U_1' \cong \partial U_1 \cup D^3$ so $V_1 \cap U_1'$ is 1-connected, and since $\partial U_1' \cong V_1 \cap U_1' - D^3$, it follows that $\partial U_1'$ is 1-connected. Since ∂D^3 is a generator x of kernel i_{1*}, $i_1\colon \partial U_1 \to V_1$, it follows that kernel $i_{1*}' \cong$ kernel $i_{1*}/(x)$, so that continuing

in this way we will finally arrive at $U_2 \supset U_1$ such that $i_{2*}: H_2(U_2) \to H_2(W)$ is an isomorphism, with ∂U_2 and $V_2 = \text{closure } W - U_2$ still 1-connected.

We will need the following (cf. [6; Lemma 1]).

LEMMA 8. *Let X be an n-manifold with boundary $\partial X = M \cup N$, M, N, X all 1-connected, $n \geqq 6$. Suppose $\pi_k(X, M) = 0$ for $2 \leqq k \leqq r - 1 < n - 4$. Then any element $w \in H_{r+1}(X, M)$ can be represented by an embedded handle $D^{r+1} \times D^{n-r-1} \subset X$, meeting ∂X normally in $S^r \times D^{n-r-1} \subset M$.*

Proof. We give the proof in the smooth case using the handle body theory of Smale and for the combinatorial situation we confine ourselves to remarking that the analogous facts are true in that case, as has been shown by Stallings.

Now a theorem of Smale [7; 6.5] says that under our hypotheses, X has a handle decomposition $X = \bigcup_{i=r-1}^{n} X_i$, where $X_{r-1} = M \times I$ and

$$X_j = X_{j-1} \cup H_1{}^j \cup \cdots \cup H_q{}^j,\ H_i{}^j = D^j \times D^{n-j},\ H_i{}^j \cap H_k{}^j = \emptyset,\quad i \neq k,$$
$$H_i{}^j \cap X_{j-1} = S^{j-1} \times D^{n-j} \subset \partial X_{j-1} - (M \times 0).$$

Since X_j is the homotopy type of X_{j-1} with some j-dimensional disks attached, it follows that $h_j: H_k(X_j) \to H_k(X)$ is an isomorphism for $k < j$ and onto for $k = j$, so that $\bar{h}_j: H_k(X_j, M) \to H_k(X, M)$ is iso for $k < j$ and onto for $k = j$. Hence there is a $w' \in H_{r+1}(X_{r+1}, M)$ such that $\bar{h}_{r+1} w' = w$.

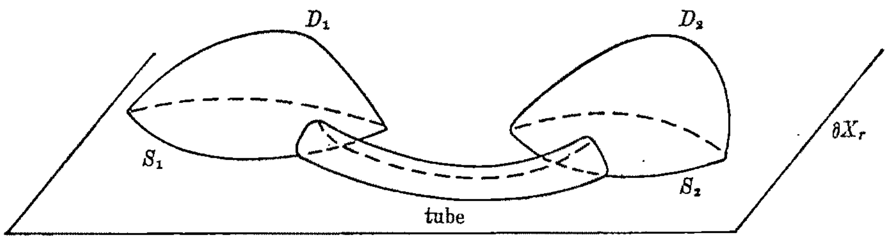
Consider the exact sequence

$$\cdots \to H_{r+1}(X_r, M) \to H_{r+1}(X_{r+1}, M) \xrightarrow{k_*} H_{r+1}(X_{r+1}, X_r)$$
$$\xrightarrow{\partial} H_r(X_r, M) \to \cdots.$$

Since $X_r \cong M \bigcup \cup D^r$, $H_{r+1}(X_r, M) = 0$, so $H_{r+1}(X_{r+1}, M)$ is mapped isomorphically by k_* onto the kernel $\partial \subset H_{r+1}(X_{r+1}, X_r)$. Let $y = k_* w'$.

Now $H_{r+1}(X_{r+1}, X_r)$ is a free abelian group with generators represented by the relative homology classes of the "core" of each handle, i.e. $(D^{r+1} \times 0, S^{r+1} \times 0) \subset (H^{r+1}, \partial H^{r+1})$. Smale [7] has shown (see also Wallace [10]) that if we are given any basis for $H_{r+1}(X_{r+1}, X_r)$ that we may find a set of handles in X_{r+1} attached to X_r so that $X_{r+1} = X_r \cup$ (these handles) and the cores of the new handles yield the given basis for $H_{r+1}(X_{r+1}, X_r)$. Hence we may assume that there is a handle such that $y = mz$, z representing the core of one of the handles of X_{r+1}. But if the core is of codimension > 1, which is the case here, any multiple of its homology is represented by a handle also. If the handle is $D^{r+1} \times D^{n-r-1}$, then we may embed in it m disjoint disks of the form $D_\alpha = D^{r+1} \times t_\alpha \subset D^{r+1} \times D^{n-r-1}$, where t_α are m disjoint points of D^{n-r-1}. Since the codimension is > 1, $\partial D_\alpha{}^{r+1}$ do not separate

$S^r \times D^{n-r-1}$, so that we may join the $S_\alpha{}^r = \partial D_\alpha{}^{r+1}$ by tubes $S^{r-1} \times I$ in $S^r \times D^{n-r-1}$ so as to form the connected sum of the $S_\alpha{}^r$, and the D_α's can be connected by tubes $D^r \times I$ in $D^{r+1} \times D^{n-r-1}$ to form the connected sum along the boundaries of the D_α, with the proper orientation. The picture gives an indication of the process:



(Two igloos connected by a tunnel might be an appropriate title.)

Then this disc and its normal bundle is $\bar{D}^{r+1} \times D^{n-r-1}$ and its core has the desired homology class $y \in H_{r+1}(X_{r+1}, X_r)$, (see [6; Lemma 1]). However, it is attached to ∂X_r rather than $M \times 1 \subset \partial X_{r-1}$, so it remains to show that it can be chosen to miss the handles of X_r and thus be attached to $M \times 1$.

Now if the attaching sphere $\bar{S}^r \times 0$ of this handle does not meet the transverse spheres $s \times S^{n-r-1}$ of the handles of index r in ∂X_r, then it may be moved off these handles down to $M \times 1$ by an isotopy of X_{r+1} so that the image would be the handle we need. For $D^r \times S^{n-r-1} - s \times S^{n-r-1}$ may be deformed by an isotopy into a neighborhood of $S^{r-1} \times S^{n-r-1}$ in $\partial X_r - S^{r-1} \times D^{n-r}$.

If the intersection number of $\bar{S}^r \times 0$ and $s \times S^{n-r-1}$ is 0 then since $r > 2$ and $n - r - 1 < n - 3$, it follows from Whitney's theorem [11] that we may change $\bar{S}^r \times 0$ by an isotopy to miss $s \times S^{n-r-1}$. But if $y \in H_{r+1}(X_{r+1}, X_r)$ is the homotopy class of the core of the handle $\bar{D}^{r+1} \times D^{n-r-1}$, then $\partial y = \sum \alpha_j h_j$, where h_j is the homology class of the r handles which generate $H_r(X_r, M)$ and α_j is the intersection number of $\bar{S}^r \times 0$ and the transverse sphere $s \times S_j{}^{n-r-1}$ of the j-th handle of X_r. Since $\partial y = 0$, all the intersection numbers are zero and there is an isotopy of X_{r+1} which takes $\bar{D}^{r+1} \times D^{n-r-1}$ into a handle attached to ∂X_r. This completes the proof of Lemma 8.

LEMMA 9. *Assume Proposition 5 for $k \leq n-3$, so that given compact C one can find $U_k \subset W$, U_k compact manifold with boundary, ∂U_k, and $V_k =$ closure $(W - U_k)$ 1-connected, $C \subset$ interior U_k and $i_{k*}: H_j(U_k) \to H_j(W)$ isomorphism for $j < k$, onto for all j. Then if $x \in (\text{kernel } i_{k*})_k$ there is U_k' containing U_k in its interior with all the above properties, such that $j_*'(x) = 0$ in $H_k(U_k')$, $j': U_k \subset U_k'$. Hence, if $y \in H_{k+1}(W, U_k)$, $\partial y = x \in H_k(U_k)$, there is $U_k' \supset U_k$ as above with $l_*'y = 0$ in $H_{k+1}(W, U_k')$,*

$$l': (W, U_k) \subset (W, U_k').$$

Proof. Given $x \in (\text{kernel } i_{k*})_k$, there is a compact set $D \supset U_k$ such that $j_* x = 0$, $j: U_k \subset D$. By Proposition 5, for $k \leqq n-3$ we can find U_k' with all the required properties and $D \subset \text{interior } U_k'$. Then since $j_* x = 0$ in $H_*(D)$, $j_*' x = i_{D*} j_* x = 0$ in $H_*(U_k')$.

Let $X =$ closure of $U_k' - U_k$, $X \subset V$ and consider the diagram ($U = U_k$, $U' = U_k'$)

$$\begin{array}{ccccccc}
H_k(\partial U) & \xleftarrow{\bar\partial} & H_{k+1}(V, \partial U) & \xrightarrow[q_*]{\cong} & H_{k+1}(W, U) & \xrightarrow{\partial} & H_k(U) \\
\downarrow{\scriptstyle h_*} & & \downarrow & & \downarrow{\scriptstyle l_*'} & & \downarrow{\scriptstyle j_*'} \\
H_k(X) & \xleftarrow{\bar\partial'} & H_{k+1}(V, X) & \xrightarrow{\cong} & H_{k+1}(W, U') & \xrightarrow{\partial'} & H_k(U').
\end{array}$$

Note that ∂, $\bar\partial$, ∂', $\bar\partial'$ are mono. Hence if $\partial y = x \in H_k(U)$ and $j_*'(x) = 0$ it follows that $k_*' y = 0$ proving the lemma.

Note that also $h_* \bar\partial q_*^{-1}(x) = 0$ in $H_k(X)$. Hence we get:

LEMMA 10. *With the hypothesis of Lemma 9, for and $z \in H_{k+1}(V, \partial U)$ there is compact manifold with boundary $X \subset V$, $\partial X = \partial U \cup N$ such that $h_*(\bar\partial z) = 0$ in $H_k(X)$.*

Now we prove Proposition 5 by induction, having proved it for $k = 3$. Let U_k be a manifold with the required properties for $k < n-3$, we would like to produce one with the properties for $k+1$. By Lemma 10 for any $x \in (\text{kernel } i_{k*})_k \subset H_k(U_k)$ we may find $y \in H_k(\partial U_k)$ with $l_* y = x$, $l: \partial U_k \subset U_k$ and a $U_k' = U_k \cup X$ so that $h_* y = 0$ in $H_k(X)$, and $y = =w$, $w \in H_{k+1}(X, \partial U_k)$. Then by Lemma 8 we can find a handle $\bar D^{k+1} \times D^{n-k+1} \subset X$ attached to ∂U_k which represents w, so that y goes to 0 in $\partial U_k \cup \bar D^{k+1} \times D^{n-k+1}$. Let $\bar U_k = U_k \cup \bar D^{k+1} \times D^{n-k-1}$. Then if $\bar\imath_k: \bar U_k \subset W$,

$$(\text{kernel } \bar\imath_{k*})_k \cong (\text{kernel } i_{k*})_k / (x)$$

and we continue until we have made the kernel 0. This completes the proof of Proposition 5.

5. Proof of Proposition 4; conclusion. By Proposition 5 we may assume that given C we can find a compact manifold $U \subset W$ with ∂U 1-connected, $V =$ closure $(W - U)$ 1-connected, $C \subset$ interior U and $i_*: H_i(U) \to H_i(W)$ an isomorphism for $i < n-3$, onto for all i. From the diagram

$$\begin{array}{ccccccccc}
0 \to & H_{k+1}(W, U) & \xrightarrow{\partial} & H_k(U) & \xrightarrow{i_*} & H_k(W) & \to 0 \\
& \cong \uparrow{\scriptstyle q_*} & & \uparrow{\scriptstyle q_*} & & \uparrow{\scriptstyle q_*} & \\
0 \to & H_{k+1}(V, \partial U) & \xrightarrow{\partial'} & H_k(\partial U) & \xrightarrow{j_*} & H_k(V) & \to 0
\end{array}$$

where $q\colon (V, \partial U) \subset (W, U)$ is an excision, we see that kernel $i_* \cong$ kernel j_*. Since ∂U is 1-connected $H_{n-2}(\partial U) \cong H^1(\partial U) = 0$, so that

$$H_{k+1}(W, U) = H_{k+1}(V, \partial U) = 0$$

for $k \neq n-3$ and $H_{n-3}(W, U)$ is free.

We can find a compact set D, $U \subset D \subset W$ such that $i_{D*}(\ker i_*) = 0$, $i_D\colon U \to D$ is inclusion. Let U' be a compact manifold with boundary with $U \cup D \subset$ interior U', and satisfying all the conditions of Proposition 5, as above with U. Then $h_*(\ker i_*) = 0$, $h\colon U \to U'$ is inclusion. It follows from the diagram

$$\begin{tikzcd}
0 \arrow[r] & H_{n-2}(W, U) \arrow[r] \arrow[d, "h_*"] & H_{n-3}(U) \arrow[r, "i_*"] \arrow[d, "h_*"] & H_{n-3}(W) \arrow[r] \arrow[d, "1"] & 0 \\
0 \arrow[r] & H_{n-2}(W, U') \arrow[r] & H_{n-3}(U') \arrow[r, "i_*'"] & H_{n-3}(W) \arrow[r] & 0
\end{tikzcd}$$

that $h_*(H_{n-3}(W, U)) = 0$. Set $V' =$ closure $(W - U')$, $M = \partial U$, $N = \partial U'$, $X =$ closure $(U' - U)$ so that $\partial X = M \cup N$; let $l_M\colon M \to X$, $l_N\colon N \to X$, $r\colon X \to V$ be inclusions. Then, since

$$q\colon (V, M) \to (W, U) \text{ and } \bar{q}\colon (V, X) \to (W, U')$$

are excisions, the diagram

$$\begin{tikzcd}
H_{n-2}(V, M) \arrow[r, "\cong"] \arrow[d, "\bar{h}_*"] & H_{n-2}(W, U) \arrow[d, "h_*"] \\
H_{n-2}(V, X) \arrow[r, "\cong"] & H_{n-2}(W, U')
\end{tikzcd}$$

shows that $\bar{h}_* = 0$. Since $H_{n-2}(V, X)$ and $H_{n-2}(V, M)$ are free, and $H_i(V, X) = H_i(V, M) = 0$ for $i \neq n-2$, this implies that $\bar{h}^* = 0$, $\bar{h}^*\colon H^{n-2}(V, X) \to H^{n-2}(V, M)$.

Consider the diagram with exact rows:

$$\begin{tikzcd}
0 & H^{n-2}(V, M) \arrow[l] & H^{n-3}(M) \arrow[l] & H^{n-3}(V) \arrow[l] & 0 \arrow[l] \\
0 & H^{n-2}(V, X) \arrow[l] \arrow[u, "\bar{h}^*"] \arrow[d, "h'^*"] & H^{n-3}(X) \arrow[l, "\delta"] \arrow[u, "l_M^*"] \arrow[d, "l_N^*"] & H^{n-3}(V) \arrow[l] \arrow[u, "1"] \arrow[d] & 0 \arrow[l] \\
0 & H^{n-2}(V', N) \arrow[l] & H^{n-3}(N) \arrow[l, "\delta'"] & H^{n-3}(V') \arrow[l] & 0. \arrow[l]
\end{tikzcd}$$

Since $\bar{h}^*=0$ and $H^{n-2}(V,X)$ is free, we can find $\alpha\colon H^{n-2}(V,X)\to H^{n-3}(X)$ such that $l_M{}^*\circ\alpha=0$ and $\delta\circ\alpha=1$ on $H^{n-2}(V,X)$. Since the inclusion $h'\colon(V',N)\to(V,X)$ is an excision,

$$\beta=l_N{}^*\circ\alpha\circ h'^{*-1}\colon H^{n-2}(V',N)\to H^{n-3}(N)$$

is defined, and $\delta'\circ\beta=1$ on $H^{n-2}(V',N)$. Further image $\beta\subset l_N{}^*(\text{kernel } l_M{}^*)$.

LEMMA 11. *$\cap\,\mu_N$ send $l_N{}^*$ (kernel $l_M{}^*)^{n-k-1}$ isomorphically onto (kernel $l_{N*})_k$.*

Proof. We have the commutative diagram with exact rows (see [2]):

$$\begin{array}{ccccccc}
\to H^{n-k-1}(X) & \xrightarrow{l^*} & H^{n-k-1}(\partial X) & \xrightarrow{\delta} & H^{n-k}(X,\partial X)\to \\
\downarrow{\cap\,\nu} & & \downarrow{\cap\,\bar\mu} & & \uparrow{\cap\,\nu} \\
\to H_{k+1}(X,\partial X) & \xrightarrow{\partial} & H_k(\partial X) & \xrightarrow{l_*} & H_k(X)\to
\end{array}$$

where $l\colon\partial X\subset X$, $\nu\in H_n(X,\partial X)$ is the fundamental class,

$H_*(\partial X)=H_*(N)+H_*(M)$; $\bar\mu=\partial\nu=\mu_N-\mu_M$, $l_*=l_{N*}-l_{M*}$, etc.

Then $(l^*H^{n-k-1}(X))\cap\bar\mu=(\text{kernel } l_*)_k$. Since $\cap\,\bar\mu\mid H^*(N)=\cap\,\mu_N$ and $\cap\,\bar\mu\mid H^*(M)=\cap\,\mu_M$, it follows that $(l^*H^{n-k-1}(X))\cap H^*(N)$ is mapped isomorphically by $\cap\,\mu_N$ onto $(\text{kernel } l_*)\cap H_k(N)$. But since $l^*=l_N{}^*-l_M{}^*$, $l^*H^{n-k-1}(X)\cap H_k(N)=l_N{}^*(\text{kernel } l_M{}^*)^{n-k-1}$, and

$$(\text{kernel } l_*)\cap H^*(N)=(\text{kernel } l_{N*})_k,$$

similarly, which proves the lemma.

It follows that $A=(\text{image }\beta)\cap\mu_N$ is a free direct summand of $H_2(N)$, contained in $(\text{kernel } l_{N*})_2$. Now it follows from van Kampen's Theorem that X is 1-connected, for $V=X\cup V'$, $X\cap V'=N$ and V, V', and N are all 1-connected. Then it follows from the Hurewicz Theorem that A consists of spherical cycles in N which are null homotopic in X. By an argument already used in § 4 we may find 3-handles $D_i{}^3\times D^{n-3}$ in X whose boundaries $S_i{}^2\times D^{n-3}\subset N$ are such that their homology classes are a basis for $A\subset H_2(N)\cong\pi_2(N)$. If we exchange these handles from X to V' (i.e. add them to V'), we will obtain new manifolds $\bar X=X-$interior of the handles, $\bar V=V'\cup$ (handles), $\bar N=\bar V\cap\bar X$, $V=\bar X\cup\bar V$. Now, since A is a free direct summand of $H_2(N)$, and $H_k(N)\cong H_k(V')$ for $k<n-3$, if $n>6$, we find easily that $H_j(\bar N)=H_j(N)$ for $j\neq 2$ or $n-3$, $H_j(\bar V)=H_j(V')$ for $j\neq 2$, and $H_2(\bar N)=H_2(N)/A=H_2(V)/r_*A=H_2(\bar V)$, and $\bar r_*\colon H_j(\bar N)\to H_j(\bar V)$ is still an isomorphism for $j<n-3$, onto for all j. For $n=6$,

the argument that $H_2(\bar{N}) = H_2(N)/A$ is slightly more difficult, but it follows easily from [5; Lemma 5.6]. Now by the same arguments which applied to $(V, \partial U)$, we have that $H^{n-2}(\bar{V}, \bar{N})$ is free, $H_i(\bar{V}, \bar{N}) = 0$ for $i \neq n-2$, and we have the exact sequence

$$0 \leftarrow H^{n-2}(\bar{V}, \bar{N}) \leftarrow H^{n-3}(\bar{N}) \leftarrow H^{n-3}(\bar{V}) \leftarrow 0.$$

By Poincaré duality $H^{n-3}(\bar{N}) \cong H^{n-3}(N)/A'$, where $A' = \beta H^{n-2}(V', N)$. But $H^{n-3}(\bar{V}) \cong H^{n-3}(V')$ $(n \geqq 6$, so $n-3 > 2)$, so that it follows that $H^{n-3}(\bar{N})$ and $H^{n-3}(\bar{V})$ are isomorphic groups. Since they are finitely generated and $H^{n-2}(\bar{V}, \bar{N})$ is free, it follows that $H^{n-2}(\bar{V}, \bar{N}) = 0$, so that $H_i(\bar{V}, \bar{N}) = 0$ for all i. It follows by excision, if $\bar{U} = U \cup \bar{X}$, $H_i(W, \bar{U}) = 0$ for all i, $\bar{U} \supset U \supset C$, so that Proposition 4 is proved.

6. The *h*-cobordism theorem.

We now proceed to prove Theorem 2.

Recall V is 1-connected at ∞, $\partial V = M_1 \cup M_2$, M_1, M_2 1-connected at ∞. M_1 and M_2 are deformation retracts of V, and $H_*(M_1)$ (and hence $H_*(M_2)$ and $H_*(V)$) is finitely generated.

By Theorem 1, M_1 and M_2 are isomorphic to interiors of $\bar{M}_1$, $\bar{M}_2$, compact manifolds with boundary. By taking a contraction of $\bar{M}_i$ which embeds $\bar{M}_i$ in $M_i =$ interior $\bar{M}_i$, we may consider $\bar{M}_i \subset M_i \subset V$. Using the product structure in a neighborhood of ∂V, we get embeddings of $\bar{M}_i \times [0,1] \subset V$, with $\bar{M}_i \times [0,1] \cap \partial V = \bar{M}_i \times 0 = \bar{M}_i$. Joining $\bar{M}_1 \times 1$ to $\bar{M}_2 \times 1$ by an arc in interior $V - \bigcup_{i=1,2} \bar{M}_i \times [0,1]$, and thickening the arc, we get a compact manifold U with $\partial V = \bar{M}_1 \cup W \cup \bar{M}_2$ where $\partial W = \partial \bar{M}_1 \cup \partial \bar{M}_2$, $\partial U \cap \partial V = \bar{M}_1 \cup \bar{M}_2$. Then using the same techniques which proved Theorem 1, we may enlarge U to $\bar{V} \subset V$, with V isomorphic to interior $\bar{V}$, adding things only along interior $W \subset \partial U$, so that $\partial \bar{V} = \bar{M}_1 \cup \bar{W} \cup \bar{M}_2$, $\partial \bar{W} = \partial \bar{M}_1 \cup \partial \bar{M}_2$, $\bar{W} \cap \partial V = \partial \bar{M}_1 \cup \partial \bar{M}_2$. From the diagram

$$\begin{array}{ccc} \bar{M}_i & \longrightarrow & \bar{V} \\ \big\downarrow & & \big\downarrow \\ M_i & \longrightarrow & V \end{array}$$

it follows that the inclusions $\bar{M}_i \rightarrow \bar{V}$ are homotopy equivalences, since the other maps are. It remains to show that $\bar{W}$ is an h-cobordism between $\partial \bar{M}_1$ and $\partial \bar{M}_2$.

Now by the Poincaré duality theorem, with two pieces of boundary (see [2]), we have

$$H^*(\bar{V}, \bar{M}_i) \cong H_*(\bar{V}, \bar{M}_i \cup \bar{W})$$

which is 0 since $\bar{M}_i \to \bar{V}$ is a homotopy equivalence. Hence if

$$H_*(\bar{M}_i) \xrightarrow{i_*} H_*(\bar{M}_i \cup \bar{W}) \xrightarrow{j_*} H_*(\bar{V}),$$

$j_* i_*$ is an isomorphism, j_* is an isomorphism, so that i_* is an isomorphism. Hence $0 = H_*(\bar{M}_i \cup \bar{W}, \bar{M}_i) \cong H_*(\bar{W}, \partial\bar{M}_i)$, by excision, and since $\bar{W}$ and $\partial\bar{M}_i$ are 1-connected, $\bar{W}$ is an h-cobordism.

The corollary follows by results of Smale [7].

PRINCETON UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY,
CAMBRIDGE UNIVERSITY,
CORNELL UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] W. Browder, "Structures on $M \times R$," *Proceedings of the Cambridge Pholosophical Society*, vol. 61 (1965), pp. 337-345.

[2] ———, "Cap products and Poincaré duality," *Proceedings of the American Mathematical Society* (to appear).

[3] M. Hirsch, "On combinatorial submanifolds of differentiable manifolds," *Commentarii Mathematici Helvetici*, vol. 36 (1961), pp. 103-111.

[4] M. C. Irwin, "Combinatorial embeddings of manifolds," *Bulletin of the American Mathematical Society*, vol. 68 (1962), pp. 25-27.

[5] M. Kervaire and J. Milnor, "Groups of homotopy spheres I," *Annals of Mathematics*, vol. 77 (1963), pp. 504-537.

[6] J. Levine, "Imbedding and istropy of spheres in manifolds," *Proceedings of the Cambridge Philosophical Society*, vol. 60 (1964), pp. 433-437.

[7] S. Smale, "On the structure of manifolds," *American Journal of Mathematics*, vol. 84 (1962), pp. 387-399.

[8] E. Specker, "Die erste Cohomologiegruppe von Überlagerungen und Homotopie-Eigenschaften dreidimensionalen Mannigfaltigkeiten," *Commentarii Mathematici Helvetici*, vol. 23 (1949), pp. 303-333.

[9] J. Stallings, "On the piecewise linear structure of euclidean space," *Proceedings of the Cambridge Philosophical Society*, vol. 58 (1962), pp. 481-488.

[10] A. H. Wallace, "Modifications and cobounding manifolds II," *Journal of Mathematics and Mechanics*, vol. 10 (1961), pp. 773-809.

[11] H. Whitney, "The self-intersections of a smooth n-manifold in $2n$-space," *Annals of Mathematics*, vol. 45 (1944), pp. 220-246.

INTERPOLATION OF ENTIRE FUNCTIONS.[1]

By Q. I. Rahman.

1. Introduction. It was discovered by Carlson that an entire function $f(z)$ satisfying

$$(1.1) \qquad |f(z)| < Me^{b|z|} \qquad (b < \pi)$$

is completely determined by its values at the points $ne^{i\phi}$. In fact ([23], p. 186; for other proofs see [14])

Theorem A. *If $f(z)$ is regular at all points inside the angle $-\alpha \leqq \theta \leqq \alpha$, where $\alpha \geqq \pi/2$; $|f(z)| < Me^{b|z|}$ where $b < \pi$, throughout this angle, $f(n) = 0$ for $n = 0, 1, 2, \cdots$; then $f(z)$ is identically zero.*

The example $f(z) = \sin \pi z$ shows that $b = \pi$ is not admissible.

The hypothesis placed on the growth of $f(z)$ implies that the indicator diagram[2] (defined only for $-\alpha \leqq \theta \leqq \alpha$) of the function has width less than 2π in the direction of the imaginary axis. Carlson's theorem is then included in the following more general result due to Pólya ([22]; [5], p. 153).

Theorem B. *If $f(z)$ is regular and of exponential type in the half plane $x \geqq 0$ and its indicator diagram is not bounded on the right by a vertical line segment of length 2π or more,[3] then $f(z) \equiv 0$ if $f(n) = 0$, $n = 0, 1, 2, \cdots$.*

If $f(z)$ is an entire function of exponential type and has zeros at all the integers, the following can be proved.

Theorem C. *If $f(z)$ is an entire function of exponential type whose indicator diagram has width at most 2π in the direction of the imaginary axis,*

---

[1] A major part of this paper is taken from the author's Doctoral dissertation at the University of London. The author wishes to take this opportunity to express his appreciation to Dr. J. G. Clunie without whose generous help this work would not have been possible. Thanks are also due to Professor W. K. Hayman for his valuable suggestions and constant encouragement.

Received August 21, 1964.

[2] For definition and properties of indicator diagram see ([5], Chapter 5).

[3] This means that the vertical line touching the right of the diagram does not intersect the latter in a segment of length 2π or more.

and does not contain two horizontal straight line segments whose perpendicular distance apart is 2π, then if $f(z)=0$ for $z=0, \pm 1, \pm 2, \cdots$,

$$f(z)=e^{\alpha z}\phi(z)\sin \pi z,$$

where $\phi(z)$ is an entire function of zero exponential type.

The above theorem is due to Valiron [25] and contains the following result of Pólya [22].

THEOREM D. *If $f(z)$ is an entire function of exponential type, $f(z)=0$ for $z=0, \pm 1, \pm 2, \cdots$, and*

$$(1.2) \qquad |f(z)|<\epsilon(r)e^{\pi r}, \qquad \epsilon(r)=O(r^p)$$

then $f(z)=P(z)\sin \pi z$, where $P(z)$ is a polynomial of degree not exceeding p.

In particular, if $\epsilon(r)=o(1)$ in Theorem D, $f(z)\equiv 0$.

The more general problem of determining the rate of growth of a function of exponential type along a line from its rate of growth along a sequence of points on the line was considered by Pólya. He remarked ([22], see formula (70), p. 606) that if $f(z)$ is an entire function satisfying (1.1) then the growth of the function along any radius $\arg z=\phi$ is determined by its growth at a special sequence of points in the sense that

$$(1.3) \qquad h(\phi)=\limsup_{r\to\infty} r^{-1}\log|f(re^{i\phi})|=\limsup_{n\to\infty} n^{-1}\log|f(ne^{i\phi})| \qquad (n=1,2,\cdots).$$

Various generalizations of (1.3) were made without much gain in "depth" (see [3], p. 224 and [4]). A remarkable advance was made by Miss Cartwright [10] in proving that an entire function satisfying (1.1) and bounded at the positive and negative integers is bounded along the whole of the real axis. If further $f(n)$ has a limit as n tends to infinity by positive integer values, then $f(x)$ has the same limit as x tends to infinity by positive values. Miss Cartwright also established the natural extension of the first of these results to functions regular and of finite order in an angle. Her methods depended on the use of Lagrange's interpolation formula and the Phragmén-Lindelöf principle. Other proofs have been given by Pfluger [21], Macintyre [19], Boas and Schaeffer [9] and Korevaar [15]. Macintyre developed a number of interpolation formulae from the "Borel-Laplace" transformation and obtained the extension of both the results of Miss Cartwright to functions of finite order in an angle. The basic formula is contained in the following ([19], p. 4)

THEOREM E. *If $f(z)$ is an entire function satisfying* (1.1) *and*

$$\limsup_{n\to\infty} |f(\pm n)|^{1/n} \leqq 1$$

then for $b < \omega < 2\pi - b$

$$f(z) = \lim_{\epsilon\to 0} \pi^{-1} \sum_{-\infty}^{\infty} \frac{\sin\omega(z-n)}{z-n} f(n) e^{-\epsilon|n|}. \tag{1.4}$$

He also proved the following lemma which enables many results to be generalized from entire functions of exponential type (for which proofs are often easier) to functions which are of exponential type in an angle.

MACINTYRE'S LEMMA. *Let $f(z)$ be regular and of exponential type in $|\arg z| \leqq \alpha \leqq \pi/2$, $h(\pm\alpha) < \pi \sin\alpha$, $h(0) \leqq 0$. Then*

$$f(z) = f_1(z) + f_2(z),$$

where $f_2(z)$ is an entire function of exponential type less than π satisfying the inequality $f_2(x) = O(1/|x|)$ for $x \to -\infty$ and

$$f_1(z) = O(r^{-1} e^{\pi r |\sin\theta|}), \qquad z = re^{i\theta},$$

uniformly for $|\theta| \leqq \beta < \alpha$.

Miss Cartwright's theorem was extended by Duffin and Schaeffer [11] by allowing a more general sequence $\{\lambda_n\}$, but keeping the same rate of growth for $f(z)$.

THEOREM OF DUFFIN AND SCHAEFFER. *If $f(z)$ is entire and satisfies*

$$h(\phi) = \limsup_{r\to\infty} r^{-1} \log |f(re^{i\phi})| \leqq a |\cos\phi| + b |\sin\phi| \tag{1.5}$$

with $b < \pi$; if $\{\lambda_n\}$ is an increasing sequence of real numbers such that $\lambda_{n+1} - \lambda_n \geqq 2\delta > 0$, $|n - \lambda_n| \leqq L$, $-\infty < n < \infty$, and $|f(\lambda_n)| \leqq M$, then $|f(x)| \leqq kM$ for $-\infty < x < \infty$, where k depends only on b, L and δ.

It is implied by a theorem of V. Bernstein ([5], p. 185) that for entire functions bounded at the sequence $\{\lambda_n\}$ condition (1.5) is equivalent to (1.1).

The following generalization of Cartwright's theorem in another direction is due to J. Korevaar [15].

THEOREM F. *Let $f(z)$ be regular and of exponential type in $x \geqq 0$. If $h(\pm\pi/2) < k\pi$ where k is an integer, and*

$$|f(n)| < C, |f'(n)| < C, \cdots, |f^{(k-1)}(n)| < C, \qquad n = 1, 2, \cdots,$$

then $f(x)$ is bounded for $x \geqq 0$.

A large number of papers have been published on this topic in recent years by Boas ([6] and [8]), Levinson ([18]; see also [17]), S. N. Bernstein [2], Levin [16], Ahiezer [1] and others. We quote the following theorem of Ahiezer [1] since it stands in close analogy with our first three theorems.

THEOREM OF AHIEZER. *Let* $\{\lambda_n\}_{-\infty}^{\infty}$ *be a sequence of complex numbers such that* $\lambda_0 = 0$, $|n - \lambda_n| \leqq L$ *and* $|\lambda_n - \lambda_m| \geqq 2\delta > 0$ *for* $m \neq n$. *If*

$$\psi(z) = \lim_{N\to\infty} (z - \lambda_0) \prod_{-N}^{N}{}' \left(1 - \frac{z}{\lambda_n}\right), \tag{1.6}$$

where the term corresponding to $n = 0$ *is omitted from* $\prod'$, $f(z)$ *is an entire function of exponential type, for which for a certain integer* q *and some* $\rho < \frac{1}{2}$:

a) $$\sum_{n=-\infty}^{\infty}{}' \frac{|f(\lambda_n)|^2}{|\psi'(\lambda\ )|^2\,|m|^{2q+2\rho}} < \infty;$$

b) $$\int_{-\infty}^{\infty} e^{-2\pi|y|}\,|f(iy)|^2 \frac{dy}{1+|y|^{2q-4L}} < \infty;$$

c) $$\int_1^{\infty} \frac{du}{u^2} \log \int_u^{\infty} e^{-\pi y}\,|f(\pm iy)| \frac{dy}{y^{q-4L+3}} = -\infty,$$

then

$$f(z) = \sum_{n=-\infty}^{\infty}{}' \frac{\psi(z) f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} \left(\frac{z}{\lambda_n}\right)^q + \frac{\psi(z)}{z}\; P(z)$$

where a dash to the sign of summation denotes that $n \neq 0$, *and* $P(z)$ *is the sum of the first* $q+1$ *terms of the Maclaurin series of the function* $f(z)\dfrac{z}{\psi(z)}$.

The method of Ahiezer depends on the use of "Fourier" transforms.

2. Statement of results:

2.1. We develop a general method for the treatment of this question. Our approach consists in integrating a certain function along a suitable contour. In the case of entire functions the contour is rectangular with sides parallel to the real and imaginary axes. We let the two sides parallel to the real axis go to infinity, when, as appears in the proofs, the integrals along these sides tend to zero. If now I_1 and I_2 denote the integrals along the lines parallel to the imaginary axis, then $I_1 - I_2$ is equal to the sum of a certain series representing the sum of the residues at the poles of the integrand lying inside the rectangular contour. One of the integrals I_1 and I_2 also tends to zero as the corresponding side goes to infinity. We obtain I_1 in terms of a uniformly convergent series. We use it to deduce an interpolation formula

from the "Mellin" transformation. Detailed account will be given with the proofs.

The same idea was used by Hardy ([14], pp. 330-332) to give a proof of Carlson's theorem. Although it was remarked by Macintyre ([19], see the second footnote on p. 1) that the Mellin transform method does not appear to be convenient for obtaining his results we observe that it is in fact capable of yielding far more general results and is even simpler. By our approach we obtain a number of new results. Besides, in order to show the strength of our method we deduce several of the known results mentioned in the introduction.

Unless otherwise stated $\{\lambda_n\}_{-\infty}^{\infty}$ will hereafter represent a sequence of real numbers such that $|n-\lambda_n|\leqq L$, $|\lambda_n-\lambda_m|\geqq 2\delta>0$ for $m\neq n$ and

$$\psi(z)=\lim_{N\to\infty}(z-\lambda_0)\prod_{-N}^{N}{}'\left(1-\frac{z}{\lambda_n}\right), \tag{2.1}$$

where the term corresponding to $n=0$ is omitted from $\prod'$.

We first prove an interpolation formula under three different sets of conditions.

THEOREM 1. *If $f(z)$ is an entire function satisfying*

$$\limsup_{\pm n\to\infty}|f(\lambda_n)|^{1/|\lambda_n|}\leqq 1, \tag{2.2}$$

and

$$\limsup_{\pm x\to\infty}\frac{1}{|x|}\log\int_{-\infty}^{\infty}(1+|y|)^{2Q}|f(x+iy)|e^{-\pi|y|}dy\leqq 0,\ \left(Q=L+\frac{L}{2\delta}\right) \tag{2.3}$$

then

$$f(z)=\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\frac{f(\lambda_n)}{z-\lambda_n}e^{-\epsilon|\lambda_n|}. \tag{2.4}$$

COROLLARY 1. *If $f(z)$ is an entire function satisfying*

$$\limsup_{\pm x\to\infty}\frac{1}{|x|}\log\int_{-\infty}^{\infty}e^{-\pi|y|}|f(x+iy)|\frac{dy}{1+|y|^{p-2Q}}\leqq 0, \tag{2.5}$$

and $f(\lambda_n)=0$ for $-\infty<n<\infty$, then $f(z)=P_1(z)\psi(z)$, where $P_1(z)$ is a polynomial of degree not exceeding $p-1$.

THEOREM 2. *If $f(z)$ is an entire function satisfying* (2.2),

$$\int_{-\infty}^{\infty}(1+|y|)^{2Q}|f(iy)|e^{-\pi|y|}dy<\infty, \tag{2.6}$$

and

$$(2.7) \qquad \int_1^\infty \frac{du}{u^2} \log \int_u^\infty \left|\frac{f(iy)}{\psi(iy)}\right| \frac{dy}{y^2} = -\infty,$$

where $\psi(z)$ is given by (2.1), *then* (2.4) *holds.*

More generally, we have

THEOREM 2′. *If $f(z)$ is an entire function satisfying* (2.2) *with $\lambda_0 = 0$, and for a certain integer q*

$$(2.8) \qquad \int_{-\infty}^\infty e^{-\pi|y|}\, |f(iy)|\, \frac{dy}{1+|y|^{q-2L+1}} < \infty,$$

$$(2.9) \qquad \int_1^\infty \frac{du}{u^2} \log \int_u^\infty \left|\frac{f(\pm iy)}{\psi(\pm iy)}\right| \frac{dy}{y^{q+3}} = -\infty,$$

then

$$(2.10) \qquad f(z) = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)}\left(\frac{z}{\lambda_n}\right)^{q+1} e^{-\epsilon|\lambda_n|} + P(z)\frac{\psi(z)}{z}$$

where dash to the sign of summation indicates that $n \neq 0$, and $P(z)$ is the sum of the first $q+2$ terms of the Maclaurin series of the function $f(z)\dfrac{z}{\psi(z)}$.

COROLLARY 2. *If $f(z)$ is an entire function satisfying* (2.2) *with $\lambda_0 = 0$, and for a certain integer q*

$$\text{b}') \qquad \int_{-\infty}^\infty e^{-2\pi|y|}\, |f(iy)|^2\, \frac{dy}{1+|y|^{2q-4L}} < \infty,$$

$$\text{c}') \qquad \int_1^\infty \frac{du}{u^2} \log \int_u^\infty \left|\frac{f(\pm iy)}{\psi(\pm iy)}\right| \frac{dy}{y^{q+3}} = -\infty,$$

then

$$(2.10) \qquad f(z) = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)}\left(\frac{z}{\lambda_n}\right)^{q} e^{-\epsilon|\lambda_n|} + \frac{\psi(z)}{z}P(z)$$

where a dash to the sign of summation indicates that $n \neq 0$, and $P(z)$ is the sum of the first $q+1$ terms of the Maclaurin series of the function $f(z)\dfrac{z}{\psi(z)}$.

It is clear that the theorem of Ahiezer can be easily deduced from the above result. In fact, conditions b′) and c′) are the same as conditions b) and c) of that theorem, whereas (2.2) is far weaker than a). If only

$$\frac{f(\lambda_n)}{\psi'(\lambda_n)\lambda_n{}^q} = O(1),$$

we may conclude by Lemma 8 that the above series is not only Abel summable but convergent.

THEOREM 3. *If $f(z)$ is an entire function satisfying* (1.1) *and* (2.2), *then* (2.4) *holds.*

This follows immediately from Theorem 1 since condition (1.1) is stronger than (2.3). The result is a generalization of that of Macintyre referred to earlier (Theorem E).

Theorem 1 is a special case of the following

THEOREM 4. *If $f(z)$ is an entire function satisfying*

$$\limsup_{\pm n\to\infty} |f^{(m)}(\lambda_n)|^{1/|\lambda_n|} \leqq 1, \qquad m = 0, 1, \cdots, k-1 \tag{2.11}$$

and for a certain integer k

$$\limsup_{\pm x\to\infty} \frac{1}{|x|} \log \int_{-\infty}^{\infty} (1+|y|)^{2kQ} |f(x+iy)| e^{-k\pi|y|} dy \leqq 0, \qquad (Q = L + \frac{L}{2\delta}) \tag{2.12}$$

then

$$f(z) = \Psi(z) \lim_{\epsilon\to 0} \sum_{-\infty}^{\infty} \frac{1}{\Psi^{(k)}(\lambda_n)} [\sum_{m=1}^{k} k(k-1)\cdots \times (k-m+1) \frac{f^{(k-m)}(\lambda_n)}{(z-\lambda_n)^m}] e^{-\epsilon|\lambda_n|}, \tag{2.13}$$

where $\Psi(z) = \{\psi(z)\}^k$.

Analogous generalizations of Theorems 2 and 3 can also be proved. In particular, if $f(z)$ is an entire function satisfying

$$|f(z)| < Me^{b|z|} \qquad (b < k\pi) \tag{2.14}$$

and for $m = 0, 1, \cdots, k-1$

$$|f^{(m)}(n)| < C, \qquad n = 0, \pm 1, \pm 2, \cdots \tag{2.15}$$

then by Littlewood's Tauberian theorem it follows from above that

$$f(z) = \sin^k \pi z \sum_{-\infty}^{\infty} \frac{1}{(\frac{d^k}{dz^k} \sin^k \pi z)_{z=n}} [\sum_{m=1}^{k} k(k-1)\cdots(k-m+1) \frac{f^{(k-m)}(n)}{(z-n)^m}].$$

If $b+\delta<k\pi$ then $f(z)\dfrac{\sin\delta(z-z_0)}{\delta(z-z_0)}$ satisfies (2.14), (2.15) and we get

$$f(z)\frac{\sin\delta(z-z_0)}{\delta(z-z_0)}=\sin^k\pi z\sum_{-\infty}^{\infty}\frac{1}{\left(\dfrac{d^k}{dz^k}\sin^k\pi z\right)_{z=n}}$$

$$\times\sum_{m=1}^{k}k(k-1)\cdots(k-m+1)\frac{\left[\dfrac{d^{k-m}}{dz^{k-m}}\left\{f(z)\dfrac{\sin\delta(z-z_0)}{\delta(z-z_0)}\right\}\right]_{z=n}}{(z-n)^m}$$

If we take $z=z_0$ and then drop the subscript we shall obtain an interpolation formula for $f(z)$ from which it will immediately follow that $f(x)=O(1)$ for $-\infty<x<\infty$. We can now deduce Theorem F by using a modified form of Macintyre's lemma.

Next, we shall indicate how it is possible to obtain the theorem of Duffin and Schaeffer by our method.

If in Theorem 3 we identify the sequence $\{\lambda_n\}_{-\infty}^{\infty}$ with the sequence of positive and negative integers we shall get a result, which, as we shall show in § 4, leads to Theorem E.

From Theorems 1 and 2 respectively we deduce the following two theorems. As compared to Miss Cartwright's result we assume less and also prove a little less.

THEOREM 5. *If $f(z)$ is an entire function satisfying*

$$\limsup_{\pm x\to\infty}\frac{1}{|x|}\log\int_{-\infty}^{\infty}|f(x+iy)|\,e^{-\pi|y|}\,dy\leqq 0, \tag{2.16}$$

and $|f(n)|<K$, $n=0,\pm1,\pm2,\cdots$, then $f(x)=O(\log|x|)$ for $-\infty<x<\infty$.

THEOREM 6. *If $f(z)$ is an entire function satisfying*

$$\int_{-\infty}^{\infty}|f(iy)|\,e^{-\pi|y|}\,dy<\infty, \tag{2.17}$$

$$\int_1^{\infty}\frac{du}{u^2}\log\int_u^{\infty}\left|\frac{f(iy)}{\sin\pi iy}\right|\frac{dy}{y^2}=-\infty, \tag{2.18}$$

and $|f(n)|<K$, $n=0,\pm1,\pm2,\cdots$, then $f(x)=O(\log|x|)$ for $-\infty<x<\infty$.

One would expect perhaps, that $f(x)=O(1)$ on the real axis.

Let $M(r)$ denote the maximum of $|f(z)|$ for $|z| = r$. Then we prove the following

THEOREM 7. *If $f(z)$ is an entire function of exponential type such that*

$$\int_0^{\infty} r^{2Q} M(r) e^{-\pi r}\, dr < \infty \qquad (Q = L + \frac{L}{2\delta}) \tag{2.19}$$

and $f(\lambda_n) = 0$ for $-\infty < n < \infty$, then $f(z) \equiv 0$.

We shall show that a special case of this result leads to Theorem D.

2.2. In this section we state results for functions regular and of exponential type in an angle. We first prove the following generalization of Macintyre's lemma.

THEOREM 8. *Let $f(z)$ be regular and of exponential type in*

$$|\arg z| \leqq \alpha \leqq \tfrac{\pi}{2},\ h(\pm\alpha) < k\pi \sin\alpha,\ h(0) < \pi(1-k).$$

Then $f(z) = f_1(z) + f_2(z)$, where $f_2(z)$ is an entire function of exponential type $< \pi$ satisfying the inequality $f_2(x) = O(\frac{1}{|x|})$ for $x \to -\infty$ and

$$f_1(z) = O(r^{-1} e^{k\pi r|\sin\theta|}) \qquad z = re^{i\theta},$$

uniformly for $|\theta| \leqq \beta < \alpha$.

Let $M(r, \alpha)$ denote the maximum of $|f(z)|$ for $|z| \leqq r$ and $|\arg z| \leqq \alpha \leqq \frac{\pi}{2}$. The example

$$f(z) = \frac{\sin \pi z}{(z+1)^2}$$

shows that if the conditions of Theorem 7 are satisfied in an angle then we cannot in general conclude that $f(z) \equiv 0$. We can however prove the following

THEOREM 9. *Let $f(z)$ be regular and of exponential type in* $|\arg z| \leqq \alpha \leqq \frac{\pi}{2}$, such that

$$\int_0^{\infty} r^{2L} M(r, \alpha) e^{-\pi r \sin\alpha}\, dr < \infty. \tag{2.20}$$

If $\{\lambda_n\}$ is an increasing sequence of positive numbers such that $\lambda_{n+1} - \lambda_n \geqq 2\delta > 0$, $|n - \lambda_n| \leqq L$, $n = 0, 1, 2, \cdots$ and $f(\lambda_n) = 0$ for $0 \leqq n < \infty$, then

$f(z) = h(z)\psi(z)$ where for any $\beta < \alpha$, $h(z)$ is regular in $|\arg z| \leqq \beta$ and tends to zero uniformly in the angle, and

$$\psi(z) = (z - \lambda_0) \lim_{N\to\infty} \prod_{n=1}^{N} (1 - \frac{z^2}{\lambda_n^2}).$$

We also prove the following

THEOREM 10. *If $f(z)$ is regular and of exponential type in $x \geqq 0$ then*

$$h(0) = \limsup_{x\to\infty} x^{-1} \log |f(x)| \leqq \max(I, k)$$

where

$$I = \limsup_{x\to\infty} x^{-1} \log \int_{-\infty}^{\infty} |f(x+iy)|\, e^{-\pi|y|}\, dy, \qquad k = \limsup_{n\to\infty} n^{-1} \log |f(n)|.$$

Following the proof of the above theorem it will be briefly indicated how a slight amendment of the proof leads to Pólya's result (1.3).

2.3. *Application.* Improving a result of Ferrar [12] on the consistency of the cardinal series, J. M. Whittaker [26] proved the following

THEOREM G. *If*

$$\sum_{-\infty}^{\infty} \frac{|a_n| \log(|n|+1)}{|n|} < \infty \tag{2.21}$$

and an entire function is defined by

$$f(x) = \sin \pi x \sum_{0}^{\infty} (-1)^n \left\{ \frac{a_n}{x-n} - \frac{a_{-n}}{x+n} \right\} \tag{2.22}$$

then provided $0 < \lambda < 1$,

$$f(x) = \tfrac{1}{\pi} \sin \pi(x-\lambda) \sum_{0}^{\infty} (-1)^n \left\{ \frac{f(n+\lambda)}{x-\lambda-n} + \frac{f(-n+\lambda)}{x-\lambda+n} \right\}. \tag{2.23}$$

A consistency theorem was later proved by Noble for cardinal series based on sequences λ_n, where

$$|\lambda_n - n| \leqq D < \frac{\log 2}{\pi}; \tag{2.24}$$

that is, series

$$f(x) = \psi(x) \sum_{-\infty}^{\infty} \frac{A_n}{(x-\lambda_n)\psi'(\lambda_n)}, \tag{2.25}$$

where $\psi(x)$ is same as defined in (2.1).

Let $\{\alpha_n\}$, $\{\beta_n\}$ $(-\infty < n < \infty)$ be real sequences such that $\{\alpha_n\}$ and $\{\beta_n - \lambda\}$ satisfy (2.24) for some real λ, and write

$$G(x) = \lim_{N\to\infty} (x-\alpha_0) \prod_{-N}^{N}{}' \left(1-\frac{x}{\alpha_n}\right), \tag{2.26}$$

$$H(x) = \lim_{M\to\infty} (x-\beta_0) \prod_{-M}^{M}{}' \left(1-\frac{x}{\beta_n}\right). \tag{2.27}$$

THEOREM OF NOBLE. *Suppose that A_n $(-\infty < n < \infty)$ is such that*

$$\sum_{-\infty}^{\infty} |A_n|^2 |n|^{4D} < \infty. \tag{2.28}$$

Then for all x we can define

$$f(x) = G(x) \sum_{-\infty}^{\infty} \frac{A_n}{(x-\alpha_n) G'(\alpha_n)}$$

and

$$H(x) \sum_{-N}^{N} \frac{f(\beta_n)}{(x-\beta_n) H'(\beta_n)} \xrightarrow[N]{} f(x).$$

Hypothesis (2.28) reduces in the case $\lambda_n \equiv n$ to

$$\sum_{-\infty}^{\infty} |A_n|^2 < \infty \tag{2.29}$$

which is more restrictive than Whittaker's condition (2.21). In § 4.3 we shall apply Theorem 1 to show that (2.29) implies a good deal more, namely

THEOREM 11. *If $\{\alpha_n\}$, $\{\beta_n\}$ are real sequences satisfying*

$$|\lambda_n - n| \leqq D < 1/24, \tag{2.30}$$

and (2.29) *holds, then we can define*

$$f(z) = G(z) \sum_{-\infty}^{\infty} \frac{A_n}{(z-\alpha_n) G'(\alpha_n)} \tag{2.31}$$

and

$$H(z) \sum_{-N}^{N} \frac{f(\beta_n)}{(z-\beta_n) H'(\beta_n)} \xrightarrow[N]{} f(z).$$

We next show that it is possible to deduce Whittaker's result (Theorem G) by our method. By the same approach we shall obtain an interpolation formula for $f(z)$ of Noble's theorem when $D < \frac{1}{4}$ and $\lambda = 0$.

THEOREM 12. *If $\{\alpha_n\}$, $\{\beta_n\}$ are real sequences satisfying*

$$|\lambda_n - n| \leqq D < \tfrac{1}{4}$$

and (2.28) *holds, then for all* z *we can define*

$$f(z)=G(z)\sum_{-\infty}^{\infty}\frac{A_n}{(z-\alpha_n)G'(\alpha_n)}$$

and

$$f(z)=H(z)\left\{\frac{f(0)}{H(0)}+\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{f(\beta_n)}{(z-\beta_n)H'(\beta_n)}\frac{z}{\beta_n}e^{-\epsilon|\beta_n|}\right\}.$$

In the range of D covered by Noble's theorem this result does not lead to his theorem.

3. Lemmas:

Lemma 1 provides certain necessary estimates for $\psi(z)$. For convenience, we suppose that $\lambda_n>0$ for $n>0$ and $\lambda_n<0$ for $n<0$; if this does not happen to be the case we re-index the λ's so that λ_0 is the one closest to 0. Since $|\lambda_0|\leqq L$ and there are at most $\frac{L}{2\delta}$ λ's between 0 and λ_0, we change the indices at most $\frac{L}{2\delta}$ and so our original conditions on λ_n are still satisfied with $Q=L+\frac{L}{2\delta}$ in place of L.

LEMMA 1. *Let* $\{\lambda_n\}_{-\infty}^{\infty}$ *be a sequence of real numbers such that* $\lambda_{n+1}-\lambda_n\geqq 2\delta>0$, $|n-\lambda_n|\leqq L$, $-\infty<n<\infty$. *If* λ_0 *is the one closest to zero, then there exist positive numbers* c_1, c_2, c_3, c_4, c'_4, c''_4, c_5, *and* c'_5 *depending only on* L *and* δ *such that*

(3.1) $\quad |\psi(z)|>c_1e^{\pi|y|}\left|\frac{y}{z^2}\right|^{2L}, \qquad |y|>L;$

(3.2) $\quad |\psi(x+iy)|>c_2|x|^{-c_3}$, *for* $|y|<L$ *and large* x *such that* $\lambda_n+\delta/2\leqq x\leqq\lambda_n+\delta;$

(3.3) $\quad |\psi(z)|>c_4e^{\pi|y|}\left|\frac{z^2}{y}\right|^{2L}, \qquad |y|>L;$

(3.3′) $\quad |\psi(z)|<c'_4e^{\pi|y|}(|z|+1)^{4L}$ *for all* z;

(3.3″) $\quad \left|\frac{\psi(z)}{z-\lambda_n}\right|<c''_4e^{\pi|y|}(|z|+1)^{4L}$ *for all* z, *where the function on the left is assumed to have its singularity at* $z=\lambda_n$ *removed.*

(3.4) $\quad |\psi'(\lambda_n)|>c_5|\lambda_n|^{-4L}$ *for* $L<\frac{1}{2}$;

(3.4′) $\quad |\psi'(\lambda_n)|>c'_5|\lambda_n|^{-4L-1}$ *in general.*

Estimates (3.1), (3.3) and (3.4′) are due to Levin [16] and (3.4) is implied by a result of Levinson ([17], see (16.04), (16.08)). In order to prove (3.2) we define

$$\psi^*(z) = \lim_{N\to\infty} (z-\mu_0) \prod_{-N}^{N}{}' (1 - z/\mu_n)$$

where

$$\begin{gathered}
\mu_0 = \lambda_0, \mu_1 = \lambda_1, \cdots, \mu_n = \lambda_n, \\
\mu_{n+1} = x = \lambda, \\
\mu_{n+2} = \lambda_{n+1}, \mu_{n+3} = \lambda_{n+2}, \cdots; \\
\mu_{-1} = \lambda_{-1}, \cdots\cdots, \mu_{-n} = \lambda_{-n}, \\
\mu_{-n-1} = \lambda_{-n} - \delta = \lambda', \\
\mu_{-n-2} = \lambda_{-n-1}, \mu_{-n-3} = \lambda_{-n-2}, \cdots.
\end{gathered}$$

Thus $\{\mu_n\}_{-\infty}^{\infty}$ is a sequence of real numbers such that $\mu_{n+1} - \mu_n \geqq \delta/2 > 0$, $|n - \mu_n| \leqq L + 1 + \frac{3}{2}\delta$, $-\infty < n < \infty$. Since

$$\psi^*(z) = \psi(z)(1 - z/\lambda)(1 - z/\lambda'),$$

we have

$$\frac{d}{dx}\psi^*(x) = -\frac{1}{\lambda}\psi(x)\left(1 + \frac{\lambda}{|\lambda'|}\right).$$

From (3.4′) it follows that there exist positive numbers c_2 and c_3 depending only on L and δ such that

$$|\psi(x)| > c_2 |x|^{-c_3}.$$

The minimum of $|\psi(x+iy)|$ evidently occurs for $y = 0$. Therefore $|\psi(x+iy)| > c_2 |x|^{-c_3}$.

For (3.3′) we simply observe that the λ_n's are real and so for fixed x, $|\psi(x+iy)|$ increases as z recedes from the real axis. Hence for $|y| \leqq 2L$, assuming $L > 0$ since the result otherwise is simple,

$$|\psi(z)| \leqq |\psi(x + i2L)| < c_4 e^{2L\pi} \left\{\frac{x^2 + (2L)^2}{2L}\right\}^{2L}$$

by (3.3). This in conjunction with (3.3) leads to (3.3′).

To prove (3.3″) consider the function

$$\frac{\psi(z)}{z - \lambda_n}$$

which is regular for $|z - \lambda_n| < \delta$ and so by the maximum modulus principle for $|z - \lambda_n| < \delta$

$$\left|\frac{\psi(z)}{z - \lambda_n}\right| \leqq 1/\delta \max_{|z-\lambda_n|=\delta} |\psi(z)|.$$

With (3.3′) this leads to (3.3″) for $|z-\lambda_n|<\delta$. When $|z-\lambda_n|\geqq\delta$ then the result comes immediately from (3.3′).

LEMMA 2. *If $f(z)$ is an entire function of exponential type and*

$$\int_{-\infty}^{\infty}(1+|y|)^{2Q}|f(x_0+iy)|e^{-\pi|y|}dy<\infty, \tag{3.5}$$

then $|y|^{2Q}|f(x_0+iy)|e^{-\pi|y|}\to 0$ *as* $\pm y\to\infty$.

The function $f(x_0+z)e^{i\pi z}(z+i)^{2Q}$ is regular and of exponential type for $y\geqq 0$. Also, from (3.5) we have

$$\int_0^{\infty}(1+y)^{2Q}|f(x_0+iy)|e^{-\pi y}dy\leqq\int_{-\infty}^{\infty}(1+|y|)^{2Q}|f(x_0+iy)|e^{-\pi|y|}dy$$
$$<\infty.$$

By a theorem of Boas ([5], p. 99) it follows that

$$(1+y)^{2Q}|f(x_0+iy)|e^{-\pi y}\to 0$$

as $y\to\infty$. Equivalently we have $y^{2Q}|f(x_0+iy)|e^{-\pi y}\to 0$ as $y\to\infty$. In the same way we can prove that $|y|^{2Q}|f(x_0+iy)|e^{-\pi|y|}\to 0$ as $y\to-\infty$.

LEMMA 3. *If $f(z)$ is regular and of exponential type for $y\geqq 0$ and*

$$\int_0^{\infty}|f(iy)|dy<\infty, \tag{3.6}$$

then there exist positive numbers K and α such that

$$\int_0^{\infty}|f(x+iy)|dy<Ke^{\alpha|x|}\int_0^{\infty}|f(iy)|dy. \tag{3.7}$$

For a proof of Lemma 3 see ([7], pp. 133-134).

From Lemma 3 we deduce

LEMMA 4. *If $f(z)$ is an entire function of exponential type and*

$$\int_0^{\infty}(1+|y|)^{2Q}|f(x+iy)|e^{-\pi|y|}dy<\infty \tag{3.8}$$

holds for some finite x then it holds for every finite x.

For definiteness we may suppose that (3.8) holds for $x=0$. If $f(z)$ is an entire function of exponential type then $f(z)e^{i\pi z}(z+i)^{2Q}$ is regular and of exponential type in $y\geqq 0$ and from (3.7) we get

$$\begin{aligned}\int_0^\infty |z|^{2Q}|f(x+iy)|e^{-\pi|y|}dy &\leqq \int_0^\infty |z+i|^{2Q}|f(x+iy)|e^{-\pi|y|}dy \\ &= O(e^{\alpha_1|x|})\int_0^\infty |1+y|^{2Q}|f(iy)|e^{-\pi|y|}dy \\ &\leqq K_1 e^{\alpha_1|x|}.\end{aligned} \tag{3.9}$$

We can similarly prove that [4]

$$\int_{-\infty}^0 |z|^{2Q}|f(x+iy)|e^{-\pi|y|}dy < K_2 e^{\alpha_2|x|}. \tag{3.10}$$

On adding the corresponding sides of (3.9) and (3.10) we get

$$\begin{aligned}\int_{-\infty}^\infty |y|^{2Q}|f(x+iy)|e^{-\pi|y|}dy &\leqq \int_{-\infty}^\infty |z|^{2Q}|f(x+iy)|e^{-\pi|y|}dy \\ &\leqq K_3 e^{\alpha_3|x|},\end{aligned}$$

where $K_3 = K_1 + K_2$ and $\alpha_3 = \max(\alpha_1, \alpha_2)$. Thus Lemma 4 is proved.

LEMMA 5. *Let* (i) $F(u) \in L(-\infty, \infty)$ *and let*

$$\int_1^\infty du/u^2 \log \int_u^\infty |F(v)|\,dv/v^2 = -\infty. \tag{ii}$$

Suppose $\mathfrak{F}(\Omega)$ *is the Fourier transform of* $F(u)$, *i.e.*

$$\mathfrak{F}(\Omega) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^\infty F(u)e^{-iu\Omega}du.$$

Let $g(z)$ *be analytic for* $(a_1 \leqq x \leqq a_2, 0 < y \leqq c)$ *and continuous in* $(a_1 \leqq x \leqq a_2, 0 \leqq y \leqq c)$. *Suppose* $\mathfrak{F}(x) = g(x)$ *for* $a_1 < b_1 \leqq x \leqq b_2 < a_2$. *Then* $\mathfrak{F}(x) = g(x)$ *for* $a_1 \leqq x \leqq a_2$.

Lemma 5 is stated and proved in ([17], p. 81).

LEMMA 6. *If*

$$\int_0^\infty r^{2Q}M(r)e^{-\pi r}dr < \infty \tag{2.19}$$

then $r^{2Q}M(r)e^{-\pi r} \to 0$ *as* $r \to \infty$.

If this is false then there exists a positive number a such that

$$\limsup_{r\to\infty} r^{2Q}M(r)e^{-\pi r} = a. \tag{3.11}$$

[4] K_1, K_2, α_1 and α_2 are positive constants.

Let $0 < b < a$; $1 < c < a/b$. From (3.11) it follows that there exists a sequence $\{r_n\}$, $r_n \to \infty$ such that for $r = r_n$, $n = 1, 2, \cdots$

$$r^{2Q} M(r) > bc\, e^{\pi r}.$$

If $r_n \leqq r \leqq r_n + \frac{1}{\pi} \log c$ then

$$r^{2Q} M(r) \geqq r_n{}^{2Q} M(r_n) > bc\, e^{\pi r_n} > b\, e^{\pi r}.$$

From the sequence $\{r_n\}$ it is possible to choose an infinite subsequence $\{r_{n_i}\}$, $r_{n_i} \to \infty$ with i such that the intervals $(r_{n_i}, r_{n_i} + \frac{1}{\pi} \log c)$, $i = 1, 2, \cdots$ are disjoint. Thus

$$\begin{aligned} \int_0^\infty r^{2Q} M(r) e^{-\pi r}\, dr &\geqq \sum_{i=1}^\infty \int_{r_{n_i}}^{r_{n_i} + 1/\pi \log c} r^{2Q} M(r) e^{-\pi r}\, dr \\ &> b \sum_{i=1}^\infty \int_{r_{n_i}}^{r_{n_i}+1/\pi \log c} dr \\ &= b/\pi \log c \sum_{i=1}^\infty 1 \\ &= \infty, \end{aligned}$$

which contradicts (2.19). Consequently $r^{2Q} M(r) e^{-\pi r} \to 0$ as $r \to \infty$.

$M(r, \alpha)$, the maximum of $|f(z)|$ for $|z| \leqq r$ and $|\arg z| \leqq \alpha \leqq \frac{\pi}{2}$, is ultimately a non-decreasing function of r when $f(z)$ is unbounded. We can therefore prove that (2.20) implies

$$\lim_{r \to \infty} r^{2Q} M(r, \alpha) e^{-\pi r \sin \alpha} = 0.$$

LEMMA 7. *If $g(t)$ is integrable in the Lebesgue sense, over $(-\pi, \pi)$ and does not vanish p.p. in this interval, then*

$$G(z) = \int_{-\pi}^{\pi} e^{zt} g(t)\, dt$$

is an entire function of exponential type such that for every θ

$$|G(-re^{i\theta})| + |G(re^{i\theta})| \neq O\{e^{cr(|\cos\theta| - \delta)}\}$$

where c is a fixed positive number and δ is any positive number.

Let A be the lower bound of numbers τ such that $g(t) = 0$ p.p. in (τ, π) and B the upper bound of numbers τ such that $g(t) = 0$ p.p. in $(-\pi, \tau)$. Since we are given $g(t) \neq 0$ in a set of positive measure it follows

that $A > B$. Also $g(t) = 0$ p.p. in (A, π) and in $(-\pi, B)$, but not in any neighborhood of A or B. Hence

$$G(z) = \int_B^A e^{zt} g(t)\, dt.$$

Put $t = \dfrac{A-B}{2\pi}\tau + \dfrac{A+B}{2}$ and we get

$$G(z) = e^{\frac{A+B}{2}z} \int_{-\pi}^{\pi} e^{\frac{A-B}{2\pi}z\tau} \left\{\frac{A-B}{2\pi}\, g\left(\frac{A-B}{2\pi}\tau + \frac{A+B}{2}\right)\right\} d\tau.$$

This leads to

$$G\left(\frac{2\pi}{A-B}z\right) = e^{\frac{(A+B)\pi}{A-B}z} \int_{-\pi}^{\pi} e^{z\tau}\tilde{g}(\tau)\, d\tau$$

where $\tilde{g}(\tau)$ does not vanish p.p. in any neighborhood of either π or $-\pi$. By a result of Titchmarsh ([24], Lemma 2.3) the integral on the right is a function of exponential type which is $\neq O\{e^{\pi r(|\cos\theta|-\delta)}\}$ for any θ where δ is any positive number. Hence

$$G(z) \neq O\{e^{\frac{A+B}{2}r\cos\theta + \frac{A-B}{2}r(|\cos\theta|-\delta)}\}$$

under same conditions on θ and δ. Now $\cos(\pi+\theta) = -\cos\theta$ and so either $\dfrac{A+B}{2}\cos\theta \geqq 0$ or $\dfrac{A+B}{2}\cos(\pi+\theta) \geqq 0$. Therefore the lemma follows with $c = \dfrac{A-B}{2} + \dfrac{|A+B|}{2}$.

We also need the following

LEMMA 8. *Let $\{\lambda_n\}_{-\infty}^{\infty}$ be a sequence of real numbers such that $\lambda_{n+1} - \lambda_n \geqq 2\delta > 0$, $|n - \lambda_n| \leqq L$, $-\infty < n < \infty$. If $a_n = O(\frac{1}{|n|})$ then the existence of*

$$\lim_{\epsilon\to 0} \sum_{-\infty}^{\infty} a_n e^{-\epsilon|\lambda_n|}$$

implies that $\sum_{-\infty}^{\infty} a_n$ converges and its sum is equal to the above limit.

Lemma 8 is a generalization of Littlewood's Tauberian Theorem. For a proof see ([27], pp. 195-196).

4. Proofs of the theorems:

4.1. *Proof of Theorem* 1. From (2.3) it follows that corresponding to each $\epsilon > 0$ there exists a positive number X such that

$$\int_{-\infty}^{\infty}(1+|y|)^{2Q}\,|f(x+iy)|\,e^{-\pi|y|}\,dy < e^{\epsilon x} \tag{4.1}$$

for $x > X$.

Suppose that w is positive, m $(< N)$ an integer, $\lambda_m < \kappa < \lambda_{m+1}$ and $X < \lambda_N < \lambda < \lambda_{N+1}$.

We shall integrate the function

$$\frac{f(z)}{\psi(z)}w^z$$

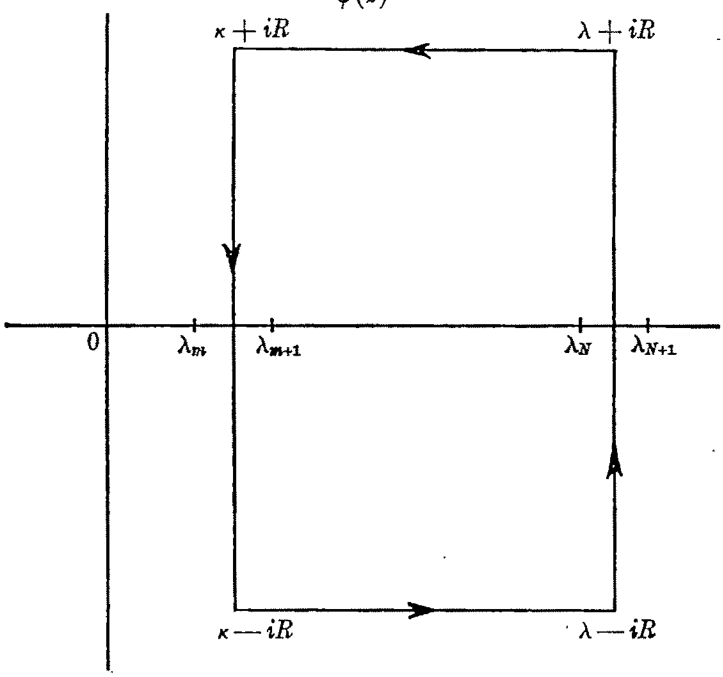


FIG. 1.

along the indicated contour. It is assumed that $w^z = e^{z \log w}$ with $\log w$ real for real positive w. If I denotes the integral over the side joining $\lambda + iR$ and $\kappa + iR$ where $R > Q$, then by (3.1)

$$|I| = \left|\int_{\lambda+iR}^{\kappa+iR}\frac{f(z)}{\psi(z)}\,w^z\,dz\right| \tag{4.2}$$

$$< 1/c_1 \int_{\kappa}^{\lambda}|f(x+iR)|\,e^{-\pi R}\left(\frac{x^2+R^2}{R}\right)^{2Q}w^x\,dx.$$

From (4.1) and Lemma 2 it follows that for a fixed x_p the function $f(x_p+z)e^{i\pi z}(z+i)^{2Q}\to 0$ as $z\to\infty$ along the imaginary axis. The function $f(x_p+z)e^{i\pi z}(z+i)^{2Q}$ is regular and of exponential type in the upper half plane. By a known theorem ([5], p. 83) it follows that $f(x+iy)e^{i\pi z}(z+i)^{2Q}\to 0$ as $y\to\infty$ for every fixed x, uniformly in any bounded x-range. Since the range of integration in (4.2) is finite we conclude that the integral I tends to zero as $R\to\infty$. We can similarly prove that the integral along the line joining $\kappa-iR$ and $\lambda-iR$ also tends to zero as $R\to\infty$. We therefore have

$$\sum_{m+1}^{N}\frac{f(\lambda_n)}{\psi'(\lambda_n)}\,w^{\lambda_n}=\frac{1}{2\pi i}\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz-\frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\frac{f(z)}{\psi(z)}\,w^z\,dz. \tag{4.3}$$

Let us suppose now that $\lambda=\lambda_N+\delta$ and that $N\to\infty$. Then

$$\begin{aligned}\left|\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz\right| &\leqq w^{\lambda_N+\delta}\int_{-\infty}^{\infty}\left|\frac{f(\lambda+iy)}{\psi(\lambda+iy)}\right|dy\\ &< w^{\lambda_N+\delta}\int_{-Q}^{Q}\left|\frac{f(\lambda+iy)}{\psi(\lambda+iy)}\right|dy\\ &\quad+\frac{w^{\lambda_N+\delta}}{c_1}\Big(\int_{-\infty}^{-Q}+\int_{Q}^{\infty}\Big)|f(\lambda+iy)|\,e^{-\pi|y|}\Big(\frac{\lambda^2+y^2}{y}\Big)^{2Q}dy\end{aligned} \tag{4.4}$$

by (3.1). Making use of (3.2) we get [5]

$$\begin{aligned}\left|\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{f(z)}{\psi(z)}\,w^z\,dz\right| &< \frac{w^{\lambda_N+\delta}}{c_2}\,(\lambda_N+\delta)^{c_3}\int_{-Q}^{Q}|f(\lambda+iy)|\,dy\\ &+c_6w^{\lambda_N+\delta}(\lambda_N+\delta)^{4Q}\Big(\int_{-\infty}^{-Q}+\int_{Q}^{\infty}\Big)|f(\lambda+iy)|\,e^{-\pi|y|}(1+|y|)^{2Q}\,dy\\ &< c_7w^{\lambda_N+\delta}(\lambda_N+\delta)^{c_3}\int_{-Q}^{Q}|f(\lambda+iy)|\,dy\\ &+c_6w^{\lambda_N+\delta}(\lambda_N+\delta)^{4Q}\Big(\int_{-\infty}^{-Q}+\int_{Q}^{\infty}\Big)|f(\lambda+iy)|\,e^{-\pi|y|}(1+|y|)^{2Q}\,dy\\ &< c_8w^{\lambda_N+\delta}\lambda_N^{c_9}\int_{-\infty}^{\infty}|f(\lambda+iy)|\,e^{-\pi|y|}(1+|y|)^{2Q}\,dy\\ &< c_{10}w^{\lambda_N+\delta}\;e^{2\epsilon(\lambda_N+\delta)}\\ &=c_{10}(e^{2\epsilon}w)^{\lambda_N+\delta}.\end{aligned}$$

Corresponding to each $w<1$ we can choose ϵ such that $e^{2\epsilon}w<1$ and then by

---

[5] c_6, c_7, etc. denote positive constants.

taking λ_N sufficiently large we can make $c_{10}(e^{2\epsilon}w)^{\lambda_N+\delta}$ arbitrarily small. Thus from (4.3) we get

$$\varnothing(w)=\frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\frac{f(z)}{\psi(z)}\,w^z\,dz-\sum_{m+1}^{\infty}\frac{f(\lambda_n)}{\psi'(\lambda_n)}\,w^{\lambda_n}\quad(0\leqq w<1). \tag{4.5}$$

By integrating

$$\frac{f(z)}{\psi(z)}\,w^z$$

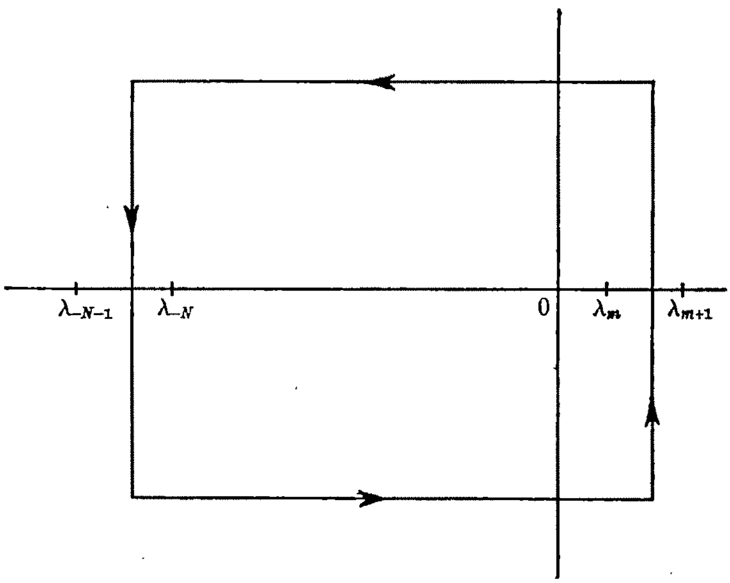


FIG. 2.

along the contour of Fig. 2 we can show that for $w\geqq 1+\epsilon>1$

$$\begin{aligned}\varnothing(w)&=\frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\frac{f(z)}{\psi(z)}\,w^z\,dz\\&=\sum_{-m}^{\infty}\frac{f(\lambda_{-n})}{\psi'(\lambda_{-n})}\,w^{\lambda_{-n}}.\end{aligned} \tag{4.6}$$

Suppose now that $-\lambda_{m+1}<\operatorname{Re} z<-\lambda_m$. Then we can choose μ and ν such that $\lambda_m<\nu<-\operatorname{Re} z<\mu<\lambda_{m+1}$. And we have

$$\int_0^1 w^{z-1}\varnothing(w)\,dw=\frac{1}{2\pi i}\int_0^1 w^{z-1}\,dw\int_{\mu-i\infty}^{\mu+i\infty}\frac{f(s)}{\psi(s)}\,w^s\,ds, \tag{4.7}$$

$$\int_1^{\infty} w^{z-1}\varnothing(w)\,dw=\frac{1}{2\pi i}\int_1^{\infty} w^{z-1}\,dw\int_{\nu-i\infty}^{\nu+i\infty}\frac{f(s)}{\psi(s)}\,w^s\,ds. \tag{4.8}$$

The double integral

$$\int_0^1 \int_{\mu-i\infty}^{\mu+i\infty} \frac{f(s)}{\psi(s)} w^{s+z-1}\, dw ds$$

is absolutely convergent as may be seen by comparison with the integral

$$\int_0^1 \int_{-\infty}^{\infty} w^{\mu+\mathrm{Re}\, z-1} \left| \frac{f(\mu+iy)}{\psi(\mu+iy)} \right| dwdy,$$

whose convergence follows from Lemma 4.

Hence the integral (4.7) is absolutely convergent and may be calculated by inversion of the order of integration. Similarly (4.8) is absolutely convergent. Inverting the order of integration and combining the results, we obtain

$$\int_0^\infty w^{z-1}\emptyset(w)\, dw = \frac{1}{2\pi i} \int_{\mu-i\infty}^{\mu+i\infty} \frac{f(s)}{\psi(s)} \frac{ds}{s+z} - \frac{1}{2\pi i} \int_{\nu-i\infty}^{\nu+i\infty} \frac{f(s)}{\psi(s)} \frac{ds}{s+z}$$

or for $-\lambda_{m+1} < \mathrm{Re}\, z < -\lambda_m$

$$\int_0^\infty w^{z-1}\emptyset(w)\, dw = \frac{f(-z)}{\psi(-z)}. \tag{4.9}$$

For $\lambda_m < \mathrm{Re}\, z < \lambda_{m+1}$ we have therefore

$$f(z) = \psi(z) \int_0^\infty w^{-z-1}\emptyset(w)\, dw$$

$$= \psi(z) \lim_{\epsilon\to 0} [\int_0^{1-\epsilon} w^{-z-1}\emptyset(w)\, dw + \int_{1+\epsilon}^\infty w^{-z-1}\emptyset(w)\, dw].$$

Replacing $\emptyset(w)$ by $-\sum_{m+1}^{\infty} \frac{f(\lambda_n)}{\psi'(\lambda_n)} w^{\lambda_n}$ in the first integral by $\sum_{-m}^{\infty} \frac{f(\lambda_{-n})}{\psi'(\lambda_{-n})} w^{\lambda_{-n}}$ in the second we get

$$f(z) = \psi(z) \lim_{\epsilon\to 0} [\int_0^{1-\epsilon} -\sum_{m+1}^{\infty} \frac{f(\lambda_n)}{\psi'(\lambda_n)} w^{\lambda_n - z-1} dw + \int_{1+\epsilon}^{\infty} \sum_{-m}^{\infty} \frac{f(\lambda_{-n})}{\psi'(\lambda_{-n})} w^{\lambda_{-n}-z-1} dw].$$

Making use of (2.2) and (3.4′) we can prove that the series on the right are uniformly convergent so that we can integrate term by term. And we get

$$f(z) = \psi(z) \lim_{\epsilon\to 0} \sum_{-\infty}^{\infty} \frac{1}{\psi'(\lambda_n)} \frac{f(\lambda_n)}{z-\lambda_n} e^{-\epsilon|\lambda_n|}. \tag{2.4}$$

The right hand side in (2.4) is continuous when $\mathrm{Re}\, z = \lambda_m$ or λ_{m+1} and so (2.4) holds for these values of z as well. In our assumption $\lambda_m \leqq \mathrm{Re}\, z \leqq \lambda_{m+1}$, m is arbitrary and therefore (2.4) holds for all z. Thus Theorem 1 is proved.

Proof of Corollary 1. Applying Theorem 1 to the function

$$1/z^p\{f(z)-\sum_{k=0}^{p-1} z^k f^{(k)}(0)\}$$

we get

$$1/z^p\{f(z)-\sum_{k=0}^{p-1} z^k f^{(k)}(0)\}=\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\frac{f(\lambda_n)-\sum_{k=0}^{p-1}\lambda_n{}^k f^{(k)}(0)}{\lambda_n{}^p(z-\lambda_n)}e^{-\epsilon|\lambda_n|}$$

$$=-\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\sum_{k=0}^{p-1}\frac{f^{(k)}(0)}{\lambda_n{}^{p-k}(z-\lambda_n)}e^{-\epsilon|\lambda_n|}.$$

Since

$$\frac{z^p}{\lambda_n{}^{p-k}(z-\lambda_n)}=z^k\Big[\frac{z^{p-k-1}}{\lambda_n{}^{p-k}}+\frac{z^{p-k-2}}{\lambda_n{}^{p-k-1}}+\cdots+\frac{1}{\lambda_n}+\frac{1}{z-\lambda_n}\Big]$$

we get

$$f(z)=-\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\sum_{k=0}^{p-1}z^k\Big[\frac{z^{p-k-1}}{\lambda_n{}^{p-k}}+\frac{z^{p-k-2}}{\lambda_n{}^{p-k-1}}+\cdots+\frac{1}{\lambda_n}\Big]f^{(k)}(0)e^{-\epsilon|\lambda_n|}$$

and the result follows.

Proof of Theorem 2. Let $\lambda_0\neq 0$ and consider

$$F(y)=\frac{f(iy)}{\psi(iy)}.$$

From (2.6) it follows that condition (i) of Lemma 5 is satisfied for $F(y)$. Now

$$\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{f(z)}{\psi(z)}w^z\,dz=\frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{f(iy)}{\psi(iy)}w^{iy}i\,dy=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(y)e^{-iy\Omega}\,dy$$

where $w=e^{-\Omega}$.

By the proof of Lemma 4 it is clear that there exist positive numbers K and α such that

$$\int_{-\infty}^{\infty}(1+|y|)^{2Q}|f(x+iy)|e^{-\pi|y|}\,dy<Ke^{\alpha|x|}. \tag{4.10}$$

Moreover, by Lemma 2, $(1+|y|)^{2Q}|f(x+iy)|e^{-\pi|y|}\to 0$ as $\pm y\to\infty$ uniformly in any bounded x-range. If we suppose for definiteness that $\lambda_0>0$ then we can deduce as for Theorem 1 that

$$\sum_{0}^{N}\frac{f(\lambda_n)}{\psi'(\lambda_n)}w^{\lambda_n}=\frac{1}{2\pi i}\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz-\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{f(z)}{\psi(z)}w^z\,dz\qquad(\lambda_N<\lambda<\lambda_{N+1}).$$

Now

$$\left|\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz\right|\leqq w^\lambda\int_{-\infty}^{\infty}\left|\frac{f(\lambda+iy)}{\psi(\lambda+iy)}\right|dy$$

$$=w^\lambda\int_{-Q}^{Q}\left|\frac{f(\lambda+iy)}{\psi(\lambda+iy)}\right|dy+w^\lambda\left(\int_{-\infty}^{-Q}+\int_{Q}^{\infty}\right)\left|\frac{f(\lambda+iy)}{\psi(\lambda+iy)}\right|dy$$

$$<Kw^\lambda\lambda^{c_{11}}e^{\alpha\lambda}$$

$$=K\left(we^{c_{11}\frac{\log\lambda}{\lambda}}e^{\alpha}\right)^\lambda$$

by (3.2) and (4.10). Therefore the integral tends to zero as $\lambda\to\infty$ if w is sufficiently small. For such values of w we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}F(y)e^{-iy\Omega}\,dy=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{f(z)}{\psi(z)}w^z\,dz=-\sum_{0}^{\infty}\frac{f(\lambda_n)}{\psi'(\lambda_n)}w^{\lambda_n}.$$

The series on the right represents an analytic function of w in $|w|\leqq 1-\epsilon$ $(\epsilon>0)$ cut along the negative real axis. By Lemma 5 it follows that

$$(4.11)\qquad \frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{f(z)}{\psi(z)}w^z\,dz=-\sum_{0}^{\infty}\frac{f(\lambda_n)}{\psi'(\lambda_n)}w^{\lambda_n}\qquad(0\leqq w\leqq 1-\epsilon).$$

But

$$\frac{1}{2\pi i}\int_{\kappa_0-i\infty}^{\kappa_0+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{f(z)}{\psi(z)}w^z\,dz$$

for $\lambda_{-1}<\kappa_0<\lambda_0$. Therefore if $\lambda_{-1}<\kappa_0<\lambda_0$ then

$$(4.12)\qquad \frac{1}{2\pi i}\int_{\kappa_0-i\infty}^{\kappa_0+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz=-\sum_{0}^{\infty}\frac{f(\lambda_n)}{\psi'(\lambda_n)}w^{\lambda_n}$$

for $0\leqq w\leqq 1-\epsilon$. If $\lambda_m<\kappa<\lambda_{m+1}$ where $m\geqq 0$ (say) then clearly

$$\frac{1}{2\pi i}\int_{\kappa-i\infty}^{\kappa+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz=\sum_{0}^{m}\frac{f(\lambda_n)}{\psi'(\lambda_n)}w^{\lambda_n}+\frac{1}{2\pi i}\int_{\kappa_0-i\infty}^{\kappa_0+i\infty}\frac{f(z)}{\psi(z)}w^z\,dz$$

$$=-\sum_{m+1}^{\infty}\frac{f(\lambda_n)}{\psi'(\lambda_n)}w^{\lambda_n},$$

for $0\leqq w\leqq 1-\epsilon$, by (4.12). In the same way we can prove (4.6). The proof can now be completed in the same way as in the case of Theorem 1.

In case $\lambda_0=0$ we may consider

$$\phi(z)=\frac{f(z)-f(0)}{z}.$$

Define the auxiliary function

$$\psi_1(z)=\frac{\psi(z)}{z}=\lim_{N\to\infty}-\frac{1}{\Lambda_0}(z-\Lambda_0)\prod_{-N}^{N}{}'\left(1-\frac{z}{\Lambda_n}\right)$$

where

$$\Lambda_1=\lambda_2,\ \Lambda_2=\lambda_3,\cdots\cdots,\ \Lambda_n=\lambda_{n+1},\cdots,$$
$$\Lambda_0=\lambda_1,$$
$$\Lambda_{-1}=\lambda_{-1},\ \Lambda_{-2}=\lambda_{-2},\cdots\cdots,\ \Lambda_{-n}=\lambda_{-n},\cdots.$$

Since $\Lambda_0\neq 0$ we have from above

$$\frac{f(z)-f(0)}{z}=\frac{\psi(z)}{z}\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}{}'\frac{\lambda_n}{\psi'(\lambda_n)}\,\frac{f(\lambda_n)-f(0)}{\lambda_n(z-\lambda_n)}\,e^{-\epsilon|\lambda_n|} \tag{4.13}$$

where dash to the sign of summation indicates $n\neq 0$. Consequently

$$f(z)-f(0)$$
$$=\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\,\frac{f(\lambda_n)}{z-\lambda_n}\,e^{-\epsilon|\lambda_n|}-\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\,\frac{f(0)}{z-\lambda_n}\,e^{-\epsilon|\lambda_n|}.$$

But clearly

$$f(0)=\psi(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\psi'(\lambda_n)}\,\frac{f(0)}{z-\lambda_n}\,e^{-\epsilon|\lambda_n|}$$

and we get the result in this case also.

We note that in Theorem 2 condition (2.6) may be replaced by

$$\int_{-\infty}^{\infty}(1+|y|)^{2Q}\,|f(x_1+iy)|\,e^{-\tau|y|}\,dy<\infty \tag{4.14}$$

for some x_1 such that $-\infty<x_1<\infty$.

To prove Theorem 2′ we consider the function

$$f^*(z)=zf(z)\,(z_0/z)^{q+2}$$

which satisfies (2.2), (4.14) and (2.7). In exactly the same way as for Theorem 2 we obtain for $0<\lambda_m<\operatorname{Re} z<\lambda_{m+1}$

$$zf(z)(z_0/z)^{q+2} = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{\lambda_n f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z_0/\lambda_n)^{q+2} e^{-\epsilon|\lambda_n|}$$

$$+ \psi(z) z_0^{q+2} \lim_{\epsilon\to 0} \int_{1+\epsilon}^{\infty} \frac{1}{\underline{|q+1}} \left[\frac{d^{q+1}}{dz^{q+1}} \left\{ \frac{zf(z)}{\psi(z)} w^z \right\}\right]_{z=0} w^{-z-1}\, dw$$

$$= \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{\lambda_n f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z_0/\lambda_n)^{q+2} e^{-\epsilon|\lambda_n|}$$

$$+ \psi(z) z_0^{q+2} \lim_{\epsilon\to 0} \int_{1+\epsilon}^{\infty} \frac{1}{\underline{|q+1}} \left[w^z \frac{d^{q+1}}{dz^{q+1}} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right.$$

$$\left. + {}^{q+1}C_1 w^z \log w \frac{d^q}{dz^q} \left\{ \frac{zf(z)}{\psi(z)} \right\} + \cdots \right]_{z=0} w^{-z-1}\, dw$$

$$= \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{\lambda_n f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z_0/\lambda_n)^{q+2} e^{-\epsilon|\lambda_n|}$$

$$+ \psi(z) z_0^{q+2} \left(\frac{1}{\underline{|q+1}} \frac{1}{z} \left[\frac{d^{q+1}}{dz^{q+1}} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} \right.$$

$$\left. + \frac{1}{\underline{|q}} \frac{1}{z^2} \left[\frac{d^q}{dz^q} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} + \cdots \right).$$

If $\lambda_m < \operatorname{Re} z < \lambda_{m+1} < 0$ then

$$zf(z)(z_0/z)^{q+2} = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{\lambda_n f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z_0/\lambda_n)^{q+2} e^{-\epsilon|\lambda_n|}$$

$$+ \psi(z) z_0^{q+2} \lim_{\epsilon\to 0} \int_0^{1-\epsilon} \frac{1}{\underline{|q+1}} \left[\frac{d^{q+1}}{dz^{q+1}} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} w^{-z-1}\, dw$$

$$= \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{\lambda_n f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z_0/\lambda_n)^{q+2} e^{-\epsilon|\lambda_n|}$$

$$+ \psi(z) z_0^{q+2} \left(\frac{1}{\underline{|q+1}} \frac{1}{z} \left[\frac{d^{q+1}}{dz^{q+1}} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} \right.$$

$$\left. + \frac{1}{\underline{|q}} \frac{1}{z^2} \left[\frac{d^q}{dz^q} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} + \cdots \right).$$

Thus for all $z \neq 0$

$$zf(z)(z_0/z)^{q+2} = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{\lambda_n f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z_0/\lambda_n)^{q+2} e^{-\epsilon|\lambda_n|}$$

$$+ \psi(z) z_0^{q+2} \left(\frac{1}{\underline{|q+1}} \frac{1}{z} \left[\frac{d^{q+1}}{dz^{q+1}} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} \right.$$

$$\left. + \frac{1}{\underline{|q}} \frac{1}{z^2} \left[\frac{d^q}{dz^q} \left\{ \frac{zf(z)}{\psi(z)} \right\} \right]_{z=0} + \cdots \right).$$

If we take $z = z_0$, drop the subscript and divide throughout by z, we shall get the result.

Proof of Corollary 2. $f(z)$ satisfies all the conditions of Theorem 2′ and therefore

$$f(z) = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z/\lambda_n)^{q+1} e^{-\epsilon|\lambda_n|} + \frac{\psi(z)}{z} P(z)$$

where $P(z)$ is the sum of the first $q+2$ terms of the Maclaurin series of the function $f(z)\dfrac{z}{\psi(z)}$. Since

$$\frac{1}{z-\lambda_n}\left(\frac{z}{\lambda_n}\right)^{q+1} = \frac{1}{z-\lambda_n}\left(\frac{z}{\lambda_n}\right)^{q} + \frac{z^q}{\lambda_n^{q+1}}$$

we get

$$f(z) = \psi(z) \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{f(\lambda_n)}{\psi'(\lambda_n)(z-\lambda_n)} (z/\lambda_n)^{q} e^{-\epsilon|\lambda_n|}$$
$$+ \psi(z) z^q \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{f(\lambda_n)}{\psi'(\lambda_n)\lambda_n^{q+1}} e^{-\epsilon|\lambda_n|} + \frac{\psi(z)}{z} P(z).$$

But b′) implies that $f(iy) = o\left(\dfrac{e^{\pi|y|}}{|y|^{q-2L}}\right)$ as $|y| \to \infty$. This is possible only if

$$z^q \lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty}{}' \frac{f(\lambda_n)}{\psi'(\lambda_n)\lambda_n^{q+1}} e^{-\epsilon|\lambda_n|} + \frac{P(z)}{z}$$

is a polynomial of degree at most $q-1$. This proves the result.

Proof of Theorem 4. The method of proof is the same as that for Theorem 1. Integrating

$$\frac{f(z)}{\Psi(z)} w^z$$

along the contour of Fig. 1 we shall obtain

$$(4.15) \qquad \Phi(w) = \frac{1}{2\pi i} \int_{\kappa-i\infty}^{\kappa+i\infty} \frac{f(z)}{\Psi(z)} w^z \, dz$$
$$= -\sum_{n=m+1}^{\infty} \frac{k}{\Psi^{(k)}(\lambda_n)} \sum_{m=1}^{k} {}^{k-1}C_{m-1} f^{(k-m)}(\lambda_n) w^{\lambda_n} (\log w)^{m-1}$$

for $0 \leqq w < 1$, whereas

$$(4.16) \qquad \Phi(w) = \frac{1}{2\pi i} \int_{\kappa-i\infty}^{\kappa+i\infty} \frac{f(z)}{\Psi(z)} w^z \, dz$$
$$= \sum_{n=-m}^{\infty} \frac{k}{\Psi^{(k)}(\lambda_{-n})} \sum_{m=1}^{k} {}^{k-1}C_{m-1} f^{(k-m)}(\lambda_{-n}) w^{\lambda_{-n}} (\log w)^{m-1}$$

for $|w| \geqq 1 + \epsilon > 1$.

In the same way as for (4.9) we can obtain

$$(4.17)\qquad f(z)=\Psi(z)\int_0^\infty w^{-z-1}\Phi(w)\,dw$$
$$=\Psi(z)\lim_{\epsilon\to 0}\Big[\int_0^{1-\epsilon} w^{-z-1}\Phi(w)\,dw+\int_{1+\epsilon}^\infty w^{-z-1}\Phi(w)\,dw\Big].$$

Replacing $\Phi(w)$ by

$$-\sum_{n=m+1}^{\infty}\frac{k}{\Psi^{(k)}(\lambda_n)}\sum_{m=1}^{k}{}^{k-1}C_{m-1}f^{(k-m)}(\lambda_n)\,w^{\lambda_n}(\log w)^{m-1}$$

in the first integral, by

$$\sum_{n=-m}^{\infty}\frac{k}{\Psi^{(k)}(\lambda_{-n})}\sum_{m=1}^{k}{}^{k-1}C_{m-1}f^{(k-m)}(\lambda_{-n})\,w^{\lambda_{-n}}(\log w)^{m-1}$$

in the second and integrating term by term we shall get the result.

We shall now show how we can deduce the theorem of Duffin and Schaeffer from Theorem 3. For this we choose $\epsilon_1>0$ and a positive integer k so that $k\epsilon_1<\pi-b$, $k>4Q$. If we set

$$\omega(z)=(\epsilon_1 z)^{-k}\sin^k(\epsilon_1 z)$$

and apply Theorem 3 to the function $f(z)\omega(z-z_0)$ we shall get

$$f(z)\omega(z-z_0)=\psi(z)\lim_{\epsilon\to 0}\sum_{n=-\infty}^{\infty}\frac{f(\lambda_n)}{z-\lambda_n}\,\frac{\omega(\lambda_n-z_0)}{\psi'(\lambda_n)}\,e^{-\epsilon|\lambda_n|}.$$

By Lemma 8 the series on the right is convergent and we get

$$f(z)\omega(z-z_0)=\psi(z)\sum_{n=-\infty}^{\infty}\frac{f(\lambda_n)}{z-\lambda_n}\,\frac{\omega(\lambda_n-z_0)}{\psi'(\lambda_n)}.$$

This is true for all z, and we may therefore take $z=z_0$ and drop the subscript; since $\omega(0)=1$,

$$f(z)=\sum_{-\infty}^{\infty}\frac{f(\lambda_n)}{z-\lambda_n}\,\frac{\omega(\lambda_n-z)}{\psi'(\lambda_n)}\,\psi(z).$$

This is the same as formula (10.5.12) of [5]. The proof can now be completed as in ([5], p. 193).

If the sequence $\{\lambda_n\}_{-\infty}^{\infty}$ is identical with the sequence $\{n\}_{-\infty}^{\infty}$, Theorem 3 takes the form:

If $f(z)$ is an entire function satisfying (1.1) and

$$\limsup_{n\to\infty}|f(\pm n)|^{1/n}\leqq 1,$$

then

$$f(z) = \frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{\sin \pi (z-n)}{z-n} f(n) e^{-\epsilon|n|}. \tag{4.18}$$

If z_0 is fixed and $|\omega| + b < \pi$ then we may apply (4.18) to

$$\sin \omega (z-z_0) f(z) \text{ and } \cos \omega (z-z_0) f(z)$$

so that

$$\sin \omega (z-z_0) f(z) = \frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{\sin \pi (z-n) \sin \omega (n-z_0)}{z-n} f(n) e^{-\epsilon|n|} \tag{4.19}$$

and

$$\cos \omega (z-z_0) f(z) = \frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{\sin \pi (z-n) \cos \omega (n-z_0)}{z-n} f(n) e^{-\epsilon|n|}. \tag{4.20}$$

If we take $z = z_0$ in (4.19) and drop the subscript we get

$$0 = \frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{(-1)^n \sin \pi z \sin \omega (z-n)}{z-n} f(n) e^{-\epsilon|n|}$$

$$= \frac{1}{\pi} \sin \pi z \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{(-1)^n \sin \omega (z-n)}{z-n} f(n) e^{-\epsilon|n|}.$$

If z is not an integer, then

$$\frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{(-1)^n \sin \omega (z-n)}{z-n} f(n) e^{-\epsilon|n|} = 0.$$

Multiplying both sides by $\cos \pi z$, we get

$$\frac{1}{\pi} \cos \pi z \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{(-1)^n \sin \omega (z-n)}{z-n} f(n) e^{-\epsilon|n|} = 0.$$

But $(-1)^n \cos \pi z$ is equal to $\cos \pi (z-n)$ and therefore

$$\frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{\cos \pi (z-n) \sin \omega (z-n)}{z-n} f(n) e^{-\epsilon|n|} = 0. \tag{4.21}$$

If we take $z = z_0$ in (4.20) and drop the subscript we get

$$\frac{1}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{\sin \pi (z-n) \cos \omega (z-n)}{z-n} f(n) e^{-\epsilon|n|} = f(z). \tag{4.22}$$

Subtracting both sides of (4.21) from the corresponding sides of (4.22), we get

$$f(z) = \lim_{\epsilon \to 0} \frac{1}{\pi} \sum_{-\infty}^{\infty} \frac{\sin(\pi - \omega)(z-n)}{z-n} f(n) e^{-\epsilon|n|} \tag{4.23}$$

if $|\omega| + b < \pi$. Putting $\pi - \omega = \varpi$ we get Theorem E.

Proof of Theorem 5. Consider the function

$$\phi(z) = \frac{f(z) - f(0)}{z}.$$

Clearly,

$$\limsup_{\pm x \to \infty} \frac{1}{|x|} \log \int_{-\infty}^{\infty} |\phi(x+iy)| e^{-\pi|y|} dy \leqq 0.$$

Besides,

$$|\phi(n)| < \frac{2k}{|n|}, \qquad n = \pm 1, \pm 2, \cdots$$

and $|\phi(0)| = |f'(0)|$. Hence by Theorem 1,

$$\phi(z) = \frac{\sin \pi z}{\pi} \lim_{\epsilon \to 0} \sum_{-\infty}^{\infty} \frac{(-1)^n}{z-n} \phi(n) e^{-\epsilon|n|}.$$

Since $\phi(n) = O(\frac{1}{|n|})$ the above series is not only Abel summable but convergent, and

$$\phi(z) = \frac{\sin \pi z}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n}{z-n} \phi(n).$$

For real x we have therefore

$$|\phi(x)| \leqq \frac{|\sin \pi x|}{|\pi x|} |f'(0)| + \frac{1}{\pi} \sum_{-\infty}^{\infty}{}' \frac{|\sin \pi(x-n)|}{|x-n|} |\phi(n)|$$

the term corresponding to $n = 0$ being omitted from $\sum'$. If $m - 1 < x < m$, then

$$|f(x)| < |f(0)| + \frac{1}{\pi} |\sin \pi x| \, |f'(0)| + k|x| \sum_{n=-\infty}^{m-2}{}' \frac{1}{|m-1-n| \, |n|}$$

$$+ k|x| \sum_{n=m+1}^{\infty} \frac{1}{(n-m)n} + \frac{1}{\pi} \frac{k|\sin \pi(x - \overline{m-1})|}{|x-m-1| \, |m-1|}$$

$$+ \frac{1}{\pi} \frac{k|\sin \pi(x-m)|}{|x-m| \, |m|}.$$

We may suppose for definiteness that m is positive. Then for $m-1<x<m$

$$
\begin{aligned}
|f(x)| &< k_1 + km\Big(\sum_{n=-\infty}^{-1} + \sum_{n=1}^{m-2}\Big)\frac{1}{|m-1-n|\,|n|} + km\sum_{n=m+1}^{\infty}\frac{1}{(n-m)n} \\
&= k_1 + km\sum_{n=1}^{\infty}\frac{1}{|m-1+n|\,|n|} + km\sum_{n=1}^{m-2}\frac{1}{|m-1-n|\,|n|} \\
&\qquad + km\sum_{n=m+1}^{\infty}\frac{1}{(n-m)n} \\
&= k_1 + k\left\{\frac{m}{1\cdot m} + \frac{m}{2(m+1)} + \cdots\right\} \\
&\qquad + k\left\{\frac{m}{1\cdot(m-2)} + \frac{m}{2\cdot(m-3)} + \cdots\right. \\
&\qquad \left. + \frac{m}{(m-3)\cdot 2} + \frac{m}{(m-2)\cdot 1}\right\} \\
&\qquad + k\left\{\frac{m}{1\cdot(m+1)} + \frac{m}{2\cdot(m+2)} + \cdots\right\} \\
&< k_1 + 2k\left\{\left(\frac{1}{1}-\frac{1}{m}\right) + \left(\frac{1}{2}-\frac{1}{m+1}\right) + \cdots\right\} \\
&\qquad + 2k\left\{\left(\frac{1}{1}+\frac{1}{m-2}\right) + \left(\frac{1}{2}+\frac{1}{m-3}\right) + \cdots\right. \\
&\qquad \left. + \left(\frac{1}{2}+\frac{1}{m-3}\right) + \left(\frac{1}{1}+\frac{1}{m-2}\right)\right\} \\
&\qquad + k\left\{\left(\frac{1}{1}-\frac{1}{m+1}\right) + \left(\frac{1}{2}-\frac{1}{m+2}\right) + \cdots\right\} \\
&< k_1 + k\left\{\left(\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{m-1}\right) + 4\left(\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{m-2}\right)\right. \\
&\qquad \left. + \left(\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{m}\right)\right\} \\
&\simeq k_2\log(m-1) \\
&< k_2\log x.
\end{aligned}
$$

The proof of Theorem 6 is similar and we therefore omit it.

It is interesting to note that the conclusion of Theorems 4 and 5 holds if $f(z)$ is an entire function of exponential type π which is $o(|z|)$ on the real axis and has $\{f(n)\}_{-\infty}^{\infty}$ bounded. In fact, $f(z)$ has the representation [25]

$$f(z) = \tfrac{1}{\pi}f'(0)\sin\pi z + f(0)\frac{\sin\pi z}{\pi z} + z/\pi\sin\pi z\sum_{n\neq 0}\frac{(-1)^n f(n)}{n(z-n)}$$

and as for the proof of Theorem 5 we can deduce that $f(x) = O(\log|x|)$ for $-\infty < x < \infty$.

Proof of Theorem 7. We again integrate the function

$$\frac{f(z)}{\psi(z)} w^z$$

along the contour of Fig. 1. If I denotes the integral over the lines joining $\lambda + iR$ and $\kappa + iR$ where $R > Q$, then by (3.1)

$$|I| = \left| \int_{\lambda+iR}^{\kappa+iR} \frac{f(z)}{\psi(z)} w^z \, dz \right|$$

$$< 1/c_1 \int_\kappa^\lambda e^{\pi\lambda} \mathrm{M}(\sqrt{R^2+\lambda^2}) e^{-\pi(R+\lambda)} \left(\frac{x^2+R^2}{R}\right)^{2Q} w^x \, dx$$

$$\leqq O(1) \int_\kappa^\lambda w^x e^{\pi\lambda} (\sqrt{R^2+\lambda^2})^{2Q} M(\sqrt{R^2+\lambda^2}) e^{-\pi\sqrt{R^2+\lambda^2}} \, dx$$

which tends to zero as $R \to \infty$ since the range of integration is finite and $(\sqrt{R^2+\lambda^2})^{2Q} M(\sqrt{R^2+\lambda^2}) e^{-\pi\sqrt{R^2+\lambda^2}} \to 0$ as $R \to \infty$. We can similarly prove that the integral along the line joining $\kappa - iR$ and $\lambda - iR$ tends to zero as $R \to \infty$. We therefore have

$$\sum_{m+1}^{N} \frac{f(\lambda_n)}{\psi'(\lambda_n)} w^{\lambda_n} = \frac{1}{2\pi i} \int_{\lambda-i\infty}^{\lambda+i\infty} \frac{f(z)}{\psi(z)} w^z \, dz - \frac{1}{2\pi i} \int_{\kappa-i\infty}^{\kappa+i\infty} \frac{f(z)}{\psi(z)} w^z \, dz.$$

Let us suppose now that $\lambda = \lambda_N + \delta$ and that $N \to \infty$. Then by (3.1) and (3.2)

$$\left| \int_{\lambda-i\infty}^{\lambda+i\infty} \frac{f(z)}{\psi(z)} w^z \, dz \right| < w^{\lambda_N+\delta} \int_{-\infty}^{\infty} \left| \frac{f(\lambda+iy)}{\psi(\lambda+iy)} \right| dy$$

$$< \frac{w^{\lambda_N+\delta}}{c_2(\lambda_N+\delta)^{-c_3}} \int_{-Q}^{Q} |f(\lambda+iy)| \, dy$$

$$+ c_{12} w^{\lambda_N+\delta} e^{\pi(\lambda_N+\delta)} \left(\int_{-\infty}^{-Q} \right.$$

$$\left. + \int_Q^\infty \right) M(\sqrt{\lambda^2+y^2}) e^{-\pi\sqrt{\lambda^2+y^2}} (\sqrt{\lambda^2+y^2})^{2Q} \, dy$$

$$< A(e^\pi w)^{\lambda_N+\delta}$$

where A is a constant. Thus the integral tends to zero if $w < e^{-\pi}$. We have therefore

$$\phi(w) = \frac{1}{2\pi i} \int_{\kappa-i\infty}^{\kappa+i\infty} \frac{f(z)}{\psi(z)} w^z \, dz = -\sum_{n=m+1}^{\infty} \frac{f(\lambda_n)}{\psi'(\lambda_n)} w^{\lambda_n} \equiv 0$$

for $w < e^{-\pi}$. We can similarly prove that $\phi(w) \equiv 0$ for $w > e^\pi$.

As for Theorem 1, we have

$$f(z)=\psi(z)\int_0^\infty w^{-z-1}\phi(w)\,dw$$

$$=\psi(z)\int_{e^{-\pi}}^{e^{\pi}} w^{-z-1}\phi(w)\,dw$$

$$=\psi(z)\int_{-\pi}^{\pi} e^{zt}\phi(e^{-t})\,dt$$

$$=\psi(z)\int_{-\pi}^{\pi} e^{zt}\chi(t)\,dt.$$

The above has been obtained on the assumption that $\lambda_m<\operatorname{Re} z<\lambda_{m+1}$. However from (2.19) it follows that $\chi(t)$ is absolutely integrable and so the above integral defines an entire function. Hence by analytic continuation the result is valid for all z.

If $\chi(t)\neq 0$ in a set of positive measure, by Lemma 7 there exists a positive number c such that for any fixed θ in $0\leqq\theta<2\pi$

$$\left|\int_{-\pi}^{\pi} e^{-tre^{i\theta}}\chi(t)\,dt\right|+\left|\int_{-\pi}^{\pi} e^{tre^{i\theta}}\chi(t)\,dt\right|>e^{cr(|\cos\theta|-\delta)} \qquad (\delta>0)$$

for a sequence of values of $r\to\infty$. Thus

$$|f(-re^{i\theta})|+|f(re^{i\theta})|>k\left(\frac{\sin\theta}{r}\right)^{2Q}e^{\pi r|\sin\theta|+cr|\cos\theta|-\delta cr}.$$

Choosing θ such that $\pi|\sin\theta|+c|\cos\theta|$ has its maximum value $\sqrt{\pi^2+c^2}$ we arrive at the conclusion that $f(z)$ is of type $>\pi$, when δ is chosen small enough. This is a contradiction and so $\chi(t)=0$ p. p. and therefore $f(z)\equiv 0$.

We shall now indicate how it is possible to deduce Theorem D from Theorem 7.

Let $\bar{f}(z)=\dfrac{f(z)}{p(z)}$ where $p(z)=z(z-1)(z-2)\cdots(z-p-1)$ and consider the function

$$g(z)=\bar{f}(z)-\frac{\sin\pi z}{\pi}\sum_0^{p+1}\frac{a_n}{z-n}$$

with $a_n=\bar{f}(n)$ for $0\leqq n\leqq p+1$. The function

$$F(z)=g(z+2)-g(z)$$

vanishes for $z=0,\pm1,\pm2,\cdots$ and $F(z)=O(r^{-2}e^{\pi r})$. By Theorem 7 it follows that $F(z)\equiv 0$ i.e., $g(z+2)\equiv g(z)$. It follows that $g(z)$ cannot have zeros other than simple zeros at $z=0,\pm1,\pm2,\cdots$, otherwise the

type of $g(z)$ would exceed π. As such $\dfrac{g(z)}{\sin\pi z}$ is an entire function and we have

$$g(z) = e^{az+b}\sin\pi z.$$

Here a must be zero because the type of $g(z)$ is $\leqq \pi$. Consequently

$$\frac{f(z)}{p(z)} = e^{b}\sin\pi z + \frac{\sin\pi z}{\pi}\sum_{0}^{p+1}\frac{a_n}{z-n}.$$

It follows that $f(z) = P(z)\sin\pi z$, where $P(z)$ is a polynomial of degree $\leqq p+2$. By the hypothesis $f(z) = O(r^{p}e^{\pi r})$ and therefore the degree of $P(z)$ is in fact $\leqq P$. The proof of Theorem D is complete.

4.2. *Proof of Theorem* 8. Let us choose $\rho > 1$ such that for the k of our theorem $k\rho < 1$, and consider the function

$$F(z) = f(\rho z).$$

$F(z)$ is regular and of exponential type $< \pi\sin\alpha$ in $|\arg z| \leqq \alpha$, and $h_F(0) = \lambda < \pi\rho(1-k)$.

Suppose that $0 \leqq w < 1$. We shall integrate the function

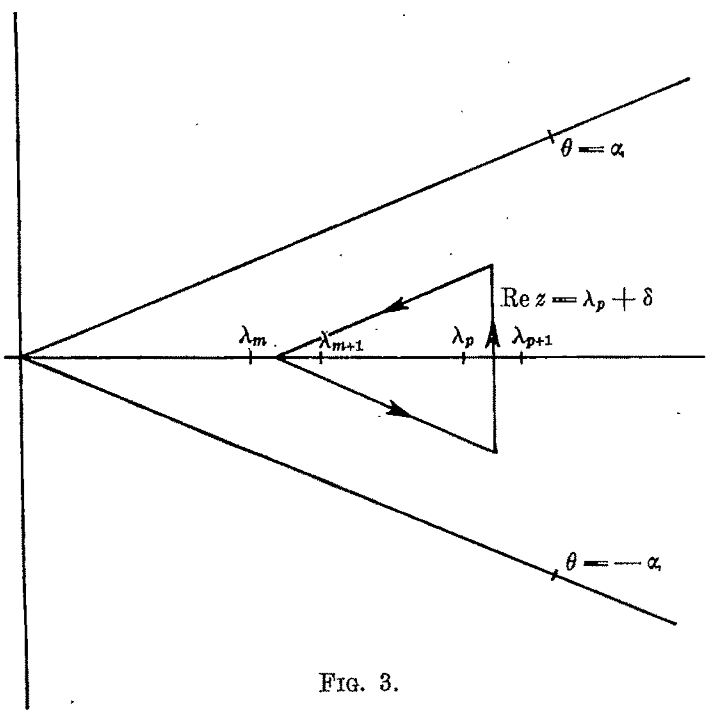


FIG. 3.

$$\frac{F(z)}{\psi(z)}w^{z}$$

along the contour of Fig. 3. Let $p \to \infty$, then for certain constants c_{13}, c_{14}, c_{15}, and c_{16} the absolute value of the integral along the side $\operatorname{Re} z = \lambda_p + \delta$ is less than or equal to

$$w^{\lambda_p+\delta} \int_{-\infty}^{\infty} \left| \frac{F(\lambda_p+\delta+iy)}{\psi(\lambda_p+\delta+iy)} \right| dy = w^{\lambda_p+\delta} \Big(\int_{|y|\leqq Q} + \int_{|y|>Q} \Big) \left| \frac{F(\lambda_p+\delta+iy)}{\psi(\lambda_p+\delta+iy)} \right| dy$$

$$< c_{13} w^{\lambda_p+\delta} (\lambda_p+\delta)^{c_3} \int_{|y|\leqq Q} | F(\lambda_p+\delta+iy) | \, dy$$

$$+ c_{14} w^{\lambda_p+\delta} (\lambda_p+\delta)^{4Q} \int_{|y|>Q} | F(\lambda_p+\delta+iy) | \, e^{-\pi|y|} (1+|y|)^{2Q} \, dy$$

$$< c_{13} w^{\lambda_p+\delta} (\lambda_p+\delta)^{c_3} e^{\pi(\lambda_p+\delta)} \int_{|y|\leqq Q} | F(\lambda_p+\delta+iy) | e^{-\pi\sqrt{(\lambda_p+\delta)^2+y^2}} \, dy$$

$$+ c_{15} w^{\lambda_p+\delta} (\lambda_p+\delta)^{c_3} e^{\pi(\lambda_p+\delta)} \int_{|y|>Q} | F(\lambda_p+\delta+iy) | e^{-\pi\sqrt{(\lambda_p+\delta)^2+y^2}} (1+|y|)^{2Q} \, dy$$

$$< c_{16} w^{\lambda_p+\delta} e^{\epsilon(\lambda_p+\delta)} \, e^{\pi(\lambda_p+\delta)}, \qquad (\epsilon > 0)$$

$$= c_{16} (e^{\epsilon} w e^{\pi})^{\lambda_p+\delta}.$$

Thus the integral tends to zero if $w < e^{-\pi}$. We have therefore

$$(4.24) \quad \phi_F(w) = \tfrac{1}{2\pi i} \int_{\kappa+\infty e^{-i\alpha}}^{\kappa} \frac{F(z)}{\psi(z)} w^z \, dz + \tfrac{1}{2\pi i} \int_{\kappa}^{\kappa+\infty e^{i\alpha}} \frac{F(z)}{\psi(z)} w^z \, dz$$

$$= - \sum_{n=m+1}^{\infty} \frac{F(\lambda_n)}{\psi'(\lambda_n)} w^{\lambda_n},$$

for $w < e^{-\pi}$. The series is convergent for $|w| \leqq e^{-\lambda-\epsilon}$, $\epsilon > 0$. Since $\phi_F(w)$ is analytic in an angular region containing $\arg w = 0$, (4.24) is valid for $0 \leqq w \leqq e^{-\lambda-\epsilon} < e^{-\lambda}$.

Suppose now that $\lambda_m < \operatorname{Re} z < \lambda_{m+1}$ and $|\arg z| \leqq \alpha$. Then we can choose μ such that $\lambda_m < \operatorname{Re} z < \mu < \lambda_{m+1}$. And we have

$$(4.25) \quad \int_0^1 w^{-z-1} \phi_F(w) \, dw = \tfrac{1}{2\pi i} \int_0^1 w^{-z-1} dw \int_{\mu+\infty e^{-i\alpha}}^{\mu} \frac{F(s)}{\psi(s)} w^s \, ds$$

$$+ \tfrac{1}{2\pi i} \int_0^1 w^{-z-1} \, dw \int_{\mu}^{\mu+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)} w^s \, ds.$$

But for certain positive constants c_{17} and c_{18}

$$\left| \int_{\mu+\infty e^{-i\alpha}}^{\mu} \frac{F(s)}{\psi(s)} w^s \, ds \right| = \left| \int_{\infty}^{0} \frac{F(\mu+re^{i\alpha})}{\psi(\mu+re^{-i\alpha})} \, w^{\mu+re^{-i\alpha}} e^{-i\alpha} \, dr \right|$$

$$< c_{17} \int_0^{\infty} | F(\mu+re^{-i\alpha}) | \, e^{-\pi r \sin\alpha} \Big(\frac{\mu^2+r^2+2\mu r\cos\alpha}{r\sin\alpha} \Big)^{2Q} w^{\mu+r\cos\alpha} \, dr$$

$$< c_{18} w^{\mu}.$$

It follows that the double integral

$$\int_0^1 w^{-z-1}\,dw \int_{\mu+\infty e^{-i\alpha}}^{\mu} \frac{F(s)}{\psi(s)}\, w^s\, ds$$

is absolutely convergent and may be calculated by inversion of the order of integration. The same argument may be applied to the integral

$$\int_0^1 w^{-z-1}\,dw \int_{\mu}^{\mu+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\, w^s\, ds.$$

Inverting the order of integration and combining the results, we obtain

$$(4.26)\qquad \int_0^1 w^{-z-1}\phi_F(w)\,dw = \frac{1}{2\pi i}\int_{\mu+\infty e^{-i\alpha}}^{\mu} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} + \frac{1}{2\pi i}\int_{\mu}^{\mu+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z}.$$

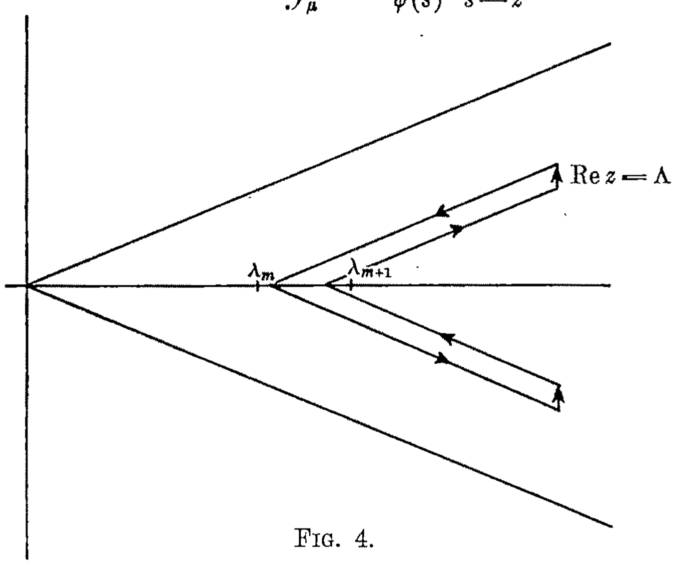


FIG. 4.

Let us choose ν such that $\lambda_m < \nu < \operatorname{Re} z < \mu < \lambda_{m+1}$ and integrate the function

$$\frac{F(s)}{\psi(s)}\,\frac{1}{s-z}$$

along the contour (enclosing z) of Fig. 4. Making $\Lambda \to \infty$ we get

$$\int_{\mu+\infty e^{-i\alpha}}^{\mu} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} + \int_{\mu}^{\mu+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} = 2\pi i\,\frac{F(z)}{\psi(z)} + \int_{\nu+\infty e^{-i\alpha}}^{\nu} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} + \int_{\nu}^{\nu+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z},$$

and (4.26) can therefore be written as

$$\int_0^1 w^{-z-1}\phi_F(w)\,dw = \frac{F(z)}{\psi(z)} + \frac{1}{2\pi i}\int_{p+\infty e^{-i\alpha}}^{p} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} + \frac{1}{2\pi i}\int_{p}^{p+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z}.$$

For z lying in the triangle Δ defined by

$$\lambda_m < \operatorname{Re} z < \lambda_{m+1}, \ |\arg(z-\lambda_m)| < \alpha < |\arg(z-\lambda_{m+1})|$$

we have therefore

$$F(z) = \psi(z)\int_0^1 w^{-z-1}\phi_B(w)\,dw - \frac{\psi(z)}{2\pi i}\int_{p+\infty e^{-i\alpha}}^{p} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} - \frac{\psi(z)}{2\pi i}\int_{p}^{p+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z}.$$

If $\infty > \delta > \lambda$ then for $z \in \Delta$ (for other values of z in $|\arg z| \leqq \alpha$ the result follows by analytic continuation)

$$\begin{aligned} F(z) &= \psi(z)\int_0^{e^{-\delta}} -w^{-z-1}\sum_{n=m+1}^{\infty}\frac{F(\lambda_n)w^{\lambda_n}}{\psi'(\lambda_n)}\,dw + \psi(z)\int_{e^{-\delta}}^1 w^{-z-1}\phi_F(w)\,dw \\ &\quad - \frac{\psi(z)}{2\pi i}\int_{p+\infty e^{-i\alpha}}^{p} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} - \frac{\psi(z)}{2\pi i}\int_{p}^{p+\infty e^{i\alpha}} \frac{F(s)}{\psi(s)}\,\frac{ds}{s-z} \\ &= \psi(z)e^{z\delta}\sum_{n=m+1}^{\infty}\frac{F(\lambda_n)e^{-\lambda_n\delta}}{\psi'(\lambda_n)(z-\lambda_n)} + \psi(z)\int_{e^{-\delta}}^1 w^{-z-1}\phi_F(w)\,dw - F_1(z) \\ &= F_2(z) + F_3(z) - F_1(z) \end{aligned}$$

where

$$F_2(z) = \psi(z)e^{z\delta}\sum_{n=m+1}^{\infty}\frac{F(\lambda_n)e^{-\lambda_n\delta}}{\psi'(\lambda_n)(z-\lambda_n)}$$

is an entire function of exponential type $\leqq \pi + \delta$. Thus

$$f(\rho z) = F_2(z) + F_3(z) - F_1(z)$$

and

$$\begin{aligned} f(z) &= F_2(z/\rho) + F_3(z/\rho) - F_1(z/\rho) \\ &= f_2(z) + f_1(z) \end{aligned}$$

where $f_2(z) = F_2(z/\rho) + F_3(z/\rho)$ is an entire function of type $\leqq (\pi+\delta)/\rho$. Since $\lambda < \pi\rho(1-k)$ it is possible to find $k' > k$ such that

$$\lambda < \pi\rho(1-k') = \lambda'.$$

Put $\delta = \lambda' = \pi\rho(1-k')$. Thus $f_2(z)$ is an entire function of exponential type $\leqq \pi/\rho + \pi(1-k')$ which can be made $< \pi$ by choosing $\rho = 1/\sqrt{kk'}$ $(k < k' < 1)$.

Let us now take for the sequence $\{\lambda_n\}_{-\infty}^{\infty}$ the sequence $\{n\}_{-\infty}^{\infty}$. $\psi(z)$ becomes $\dfrac{\sin \pi z}{\pi}$ and it is easy to verify that $f_2(z)$ so obtained satisfies $f_2(x) = O(\dfrac{1}{|x|})$ as $x \to -\infty$ and for every $k' > k$

$$f_1(z) = O(r^{-1}e^{k'\pi r|\sin\theta|}), \qquad z = re^{i\theta},$$

uniformly for $|\theta| \leqq \beta < \alpha$. We now observe that in the statement of the theorem $h(\pm\alpha)$, $h(0)$ satisfy strict inequalities and therefore we can find $\mathfrak{k}$ such that $h(\pm\alpha) < \mathfrak{k}\pi \sin\alpha < k\pi\sin\alpha$, $h(0) < \pi(1-\mathfrak{k})$. If we work with $\mathfrak{k}$ instead of k we shall get the desired result.

Proof of Theorem 9. We integrate the function

$$\frac{f(z)}{\psi(z)}\, w^z \qquad (\psi(z) = \lim_{N\to\infty}(z-\lambda_0)\prod_1^N (1 - z^2/\lambda_n^2)),$$

along the contour of Fig. 3. Let $p \to \infty$, then the absolute value of the integral along the side $\operatorname{Re} z = \lambda_p + \delta$ is for a certain constant c_{19}, less than or equal to

$$\begin{aligned} w^{\lambda_p+\delta}\int_{-\infty}^{\infty}\left|\frac{f(\lambda_p+\delta+iy)}{\psi(\lambda_p+\delta+iy)}\right| dy &< \frac{w^{\lambda_p+\delta}}{c_2(\lambda_p+\delta)^{-a_2}}\int_{-L}^{L}|f(\lambda_p+\delta+iy)|\,dy \\ &+ c_{19}w^{\lambda_p+\delta}e^{\pi(\lambda_p+\delta)}\left(\int_{-\infty}^{-L}+\int_{L}^{\infty}\right)M(\sqrt{(\lambda_p+\delta)^2+y^2},\alpha)\,e^{-\pi\sqrt{(\lambda_p+\delta)^2+y^2}} \\ &\qquad \times(\sqrt{(\lambda_p+\delta)^2+y^2})^{2L}\,dy \\ &< A'(e^{\pi}w)^{\lambda_p+\delta} \end{aligned}$$

where A' is a constant. Thus the integral tends to zero if $w < e^{-\pi}$. We have therefore

$$\begin{aligned} \phi_f(w) &= \frac{1}{2\pi i}\int_{\kappa+\infty e^{-i\alpha}}^{\kappa}\frac{f(z)}{\psi(z)}\,w^z\,dz + \frac{1}{2\pi i}\int_{\kappa}^{\kappa+\infty e^{i\alpha}}\frac{f(z)}{\psi(z)}\,w^z\,dz \\ &= -\sum_{n=m+1}^{\infty}\frac{f(\lambda_n)}{\psi'(\lambda_n)}\,w^{\lambda_n} \equiv 0 \end{aligned}$$

for $w < e^{-\pi}$.

As in the proof of Theorem 8, we have

$$\begin{aligned} f(z) = \psi(z)\int_0^1 w^{-z-1}\phi_f(w)\,dw &- \frac{\psi(z)}{2\pi i}\int_{\nu+\infty e^{-i\alpha}}^{\nu}\frac{f(s)}{\psi(s)}\,\frac{ds}{s-z} \\ &- \frac{\psi(z)}{2\pi i}\int_{\nu}^{\nu+\infty e^{i\alpha}}\frac{f(s)}{\psi(s)}\,\frac{ds}{s-z} \end{aligned}$$

$$= \psi(z) \int_{e^{-\pi}}^{1} w^{-z-1}\phi_f(w)\,dw - \frac{\psi(z)}{2\pi i}\Big(\int_{\nu+\infty e^{-i\alpha}}^{\nu} + \int_{\nu}^{\nu+\infty e^{i\alpha}}\Big)\frac{f(s)}{\psi(s)}\,\frac{ds}{s-z}$$

$$= \psi(z) \int_{-\pi}^{0} e^{zt}\phi_f(e^{-t})\,dt - \frac{\psi(z)}{2\pi i}\Big(\int_{\nu+\infty e^{-i\alpha}}^{\nu} + \int_{\nu}^{\nu+\infty e^{i\alpha}}\Big)\frac{f(s)}{\psi(s)}\,\frac{ds}{s-z}$$

$$= \psi(z) \int_{-\pi}^{0} e^{zt}\chi_1(t)\,dt - \frac{\psi(z)}{2\pi i}\Big(\int_{\nu+\infty e^{-i\alpha}}^{\nu} + \int_{\nu}^{\nu+\infty e^{i\alpha}}\Big)\frac{f(s)}{\psi(s)}\,\frac{ds}{s-z}.$$

It is easy to see that each of the two functions

$$\int_{-\pi}^{0} e^{zt}\chi_1(t)\,dt$$

and

$$\Big(\int_{\nu+\infty e^{-i\alpha}}^{\nu} + \int_{\nu}^{\nu+\infty e^{i\alpha}}\Big)\frac{f(s)}{\psi(s)}\,\frac{ds}{s-z}$$

tend to zero uniformly in the angle $|\arg z| \leqq \beta < \alpha$ and so the result follows.

Proof of Theorem 10. If $c = \max(I, k)$ the function $f^*(z) = e^{-cz}f(z)$ satisfies

$$\limsup_{n\to\infty} |f^*(n)|^{1/n} \leqq 1$$

and

$$\limsup_{x\to\infty} 1/x \log \int_{-\infty}^{\infty} |f^*(x+iy)|\, e^{-\pi|y|}\,dy \leqq 0.$$

As in the case of Theorem 1 we can prove that for $m < \mu^* < m+1$ and each $w < 1$

$$\phi^*(w) = \frac{1}{2\pi i}\int_{\mu^*-i\infty}^{\mu^*+i\infty} \frac{f^*(z)}{\sin \pi z}\, w^z\,dz = -\sum_{n=m+1}^{\infty}(-1)^n f^*(n)\,w^n.$$

Plainly we have

$$\int_0^1 w^{-s-1}\phi^*(w)\,dw = \frac{1}{2\pi i}\int_0^1 w^{-s-1}\,dw \int_{\mu^*-i\infty}^{\mu^*+i\infty} \frac{f^*(z)}{\sin \pi z}\, w^z\,dz.$$

Inverting the order of integration we have

$$(4.27) \qquad \int_0^1 w^{-s-1}\phi^*(w)\,dw = \frac{1}{2\pi i}\int_{\mu^*-i\infty}^{\mu^*+i\infty} \frac{f^*(z)}{\sin \pi z}\,\frac{dz}{s-z}.$$

Let us now choose ν^* such that $m < \nu^* < \operatorname{Re} s < \mu^* < m+1$ and integrate the function

$$\frac{f^*(z)}{\sin \pi z}\,\frac{1}{z-s}$$

along the contour of Fig. 5. The integrals along the sides parallel to the real axis tend to zero as these sides go to infinity and we get

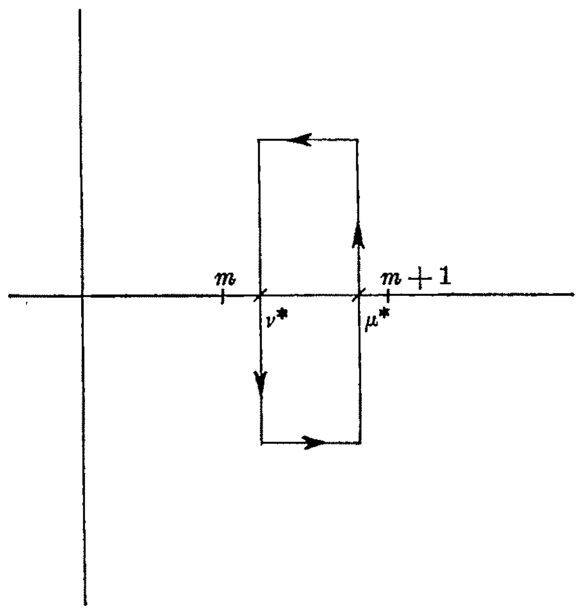


FIG. 5.

$$\int_{\mu^*-i\infty}^{\mu^*+i\infty} \frac{f^*(z)}{\sin \pi z} \frac{dz}{z-s} = 2\pi i \frac{f^*(s)}{\sin \pi s} + \int_{\nu^*-i\infty}^{\nu^*+i\infty} \frac{f^*(z)}{\sin \pi z} \frac{dz}{z-s}$$

and (4.27) can therefore be written as

$$f^*(z) = \sin \pi z \int_0^1 w^{-z-1} \phi^*(w)\, dw - \frac{\sin \pi z}{2\pi i} \int_{\nu^*-i\infty}^{\nu^*+i\infty} \frac{f^*(s)}{\sin \pi s} \frac{ds}{s-z}.$$

From this it follows that

$$\limsup_{x\to\infty} x^{-1} \log |f^*(x)| \leqq 0.$$

Since $h(0) = c + \limsup_{x\to\infty} x^{-1} \log |f^*(x)|$ we get the desired result.

To show that Pólya's theorem follows assume his conditions are satisfied with $h(0) = 0$, but $\limsup_{n\to\infty} n^{-1} \log |f(n)| = -\delta$ $(\delta > 0)$. We now find that

$$\sum_{m+1}^{\infty} \frac{(-1)^n f(n)}{\pi} w^n$$

converges for $|w| < e^{\delta}$. The integral for $\phi(w)$ is analytic in a sector, so that series and integral coincide for $0 \leqq w < e^{\delta}$. Consider

$$\int_0^{e^{\delta}} w^{-s-1}\phi(w)\,dw$$

to get.

$$f(z) = \frac{\sin \pi z}{\pi} \sum_{m+1}^{\infty} \frac{(-1)^n f(n)}{z-n} e^{\delta(n-z)} - e^{-\delta z} \frac{\sin \pi z}{2\pi i} \int_{\nu-i\infty}^{\nu+i\infty} \frac{f(s)}{\sin \pi s} \frac{e^{\delta s}}{s-z}\, ds.$$

This leads to $h(0) \leqq -\delta$ which is a contradiction.

4.3. *Proof of Theorem* 11. A dash to the sign of summation will represent that the term corresponding to α_n closest to x is omitted. Since $D < 1/24$, α_0 is the α_n closest to zero and therefore by (3.3′), (3.3″) and (3.4) we have

$$|f(x)| = |G(x) \sum_{-\infty}^{\infty} \frac{A_n}{(x-\alpha_n)G'(\alpha_n)}|$$
$$= O(|x|^{4D}) \sum_{-\infty}^{\infty}{}' \frac{|A_n|\,|\alpha_n|^{4D}}{|x-\alpha_n|} + O(|x|^{8D}).$$

By Schwarz's inequality

$$|f(x)| \leqq O(|x|^{4D}) (\sum_{-\infty}^{\infty} |A_n|^2 \sum_{-\infty}^{\infty}{}' \frac{|\alpha_n|^{8D}}{|x-\alpha_n|^2})^{1/2} + O(|x|^{8D})$$
$$= O(|x|^{4D}) \{ (\sum_{|n|<[10|x|]+2}' + \sum_{|n|\geqq[10|x|]+2}) \frac{|\alpha_n|^{8D}}{|x-\alpha_n|^2} \}^{1/2} + O(|x|^{8D})$$
$$= O(|x|^{4D}) \{ O(|x|^{8D}) + \frac{100}{81} \sum_{|n|\geqq 2} \frac{|\alpha_n|^{8D}}{|\alpha_n|^2} \}^{1/2} + O(|x|^{8D}).$$

Since $D < 1/24$, the series $\sum_{-\infty}^{\infty} \frac{|\alpha_n|^{8D}}{|\alpha_n|^2}$ converges and therefore

$$f(x) = O(|x|^{8D}). \tag{4.28}$$

Let now

$$F(z) = \frac{f(z) - f(0)}{z}.$$

From (4.28) it follows that $F(z)$ satisfies (2.2). Besides

$$\int_{-\infty}^{\infty} (1+|y|)^{2D} |F(x+iy)|\, e^{-\pi|y|}\, dy$$
$$= O(1) \int_1^{\infty} (1+y)^{2D} \frac{|f(x+iy)|}{|x+iy|} e^{-\pi y}\, dy$$
$$= O(1) \int_1^{\infty} (1+y)^{2D} \frac{|x+iy|^{4D}}{y^{2D}|x+iy|} \sum_{-\infty}^{\infty} \frac{|A_n|\,|\alpha_n|^{4D}}{\sqrt{y^2+(x-\alpha_n)^2}}\, dy$$

$$=O(|x|^{4D})\int_1^\infty \sum_{-\infty}^{\infty}{}' \frac{|A_n||\alpha_n|^{4D}y^{4D}}{|z||y||x-\alpha_n|}\,dy+O(|x|^{8D})$$

$$=O(|x|^{4D})\int_1^\infty \sum_{-\infty}^{\infty}{}' \frac{|A_n||\alpha_n|^{4D}}{y^{2-4D}|x-\alpha_n|}\,dy+O(|x|^{8D})$$

$$=O(|x|^{4D})\sum_{-\infty}^{\infty}{}' \frac{|A_n||\alpha_n|^{4D}}{|x-\alpha_n|}+O(|x|^{8D})$$

$$=O(|x|^{8D}).$$

Thus $F(z)$ also satisfies (2.3). Therefore from Theorem 1, we have [6]

[6] If $\beta_0=0$ then $\dfrac{f(\beta_0)-f(0)}{\beta_0}$ has to be replaced by $f'(0)$.

$$\frac{f(z)-f(0)}{z}=H(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{H'(\beta_n)}\,\frac{f(\beta_n)-f(0)}{\beta_n(z-\beta_n)}\,e^{-\epsilon|\beta_n|}.$$

Since $\dfrac{f(\beta_n)-f(0)}{H'(\beta_n)\beta_n(z-\beta_n)}=O\left(\dfrac{1}{|n|}\right)$, we have by Lemma 8

$$\frac{f(z)-f(0)}{z}=H(z)\left[\sum_{-N}^{N}\frac{1}{H'(\beta_n)}\,\frac{f(\beta_n)-f(0)}{\beta_n(z-\beta_n)}+\delta_N\right]$$

where $\delta_N\to 0$ as $N\to\infty$. Since

$$\frac{z}{\beta_n(z-\beta_n)}=\frac{1}{\beta_n}+\frac{1}{z-\beta_n}$$

we have

$$f(z)-H(z)\sum_{-N}^{N}\frac{1}{H'(\beta_n)}\,\frac{f(\beta_n)}{z-\beta_n}=f(0)-H(z)\sum_{-N}^{N}\frac{1}{H'(\beta_n)}\,\frac{f(\beta_n)}{\beta_n}$$

$$-zH(z)\sum_{-N}^{N}\frac{1}{\beta_nH'(\beta_n)}\,\frac{f(0)}{z-\beta_n}+zH(z)\delta_N.$$

As $N\to\infty$, $zH(z)\displaystyle\sum_{-N}^{N}\frac{1}{\beta_nH'(\beta_n)}\,\frac{f(0)}{z-\beta_n}\to f(0)-H(z)\frac{f(0)}{H(0)}$. Consequently

$$f(z)-H(z)\lim_{N\to\infty}\sum_{-N}^{N}\frac{1}{H'(\beta_n)}\,\frac{f(\beta_n)}{z-\beta_n}=H(z)\,\frac{f(0)}{H(0)}$$

$$-H(z)\lim_{N\to\infty}\sum_{-N}^{N}\frac{1}{H'(\beta_n)}\,\frac{f(\beta_n)}{\beta_n}.$$

But

$$\sum_{-N}^{N}\frac{|f(\beta_n)|}{|H'(\beta_n)||\beta_n|}\leqq\sum_{n=-N}^{N}\frac{|G(\beta_n)|}{|H'(\beta_n)||\beta_n|}\sum_{m=-\infty}^{\infty}\frac{|A_m|}{|\beta_n-\alpha_m||G'(\alpha_m)|}$$

$$=O(1)\sum_{m=-\infty}^{\infty}|A_m||\alpha_m|^{4D}\sum_{\substack{n=-N\\ n\neq m}}^{N}\frac{|\beta_n|^{4D}|\beta_n|^{4D}}{|\beta_n||\beta_n-\alpha_m|}+O(1)\sum_{-N}^{N}\frac{|A_n||\beta_n|^{8D}}{|H'(\beta_n)||\beta_n|}$$

and

$$\sum_{\substack{n=-N\\n\neq m}}^{N} \frac{1}{|\beta_n|^{1-8D}|\beta_n-\alpha_m|} = \Big(\sum_{\substack{|n|<2|m|\\n\neq m}} + \sum_{2|m|\leqq|n|\leqq N}\Big) \frac{1}{|\beta_n|^{1-8D}|\beta_n-\alpha_m|}$$

$$< \frac{1}{|\alpha_m|} \sum_{\substack{|n|<2|m|\\n\neq m}} \frac{|\beta_n|+|\beta_n-\alpha_m|}{|\beta_n|^{1-8D}|\beta_n-\alpha_m|} + O(1)\sum_{m}^{\infty} \frac{1}{|\beta_n|^{2-8D}}$$

$$< \frac{1}{|\alpha_m|} \sum_{\substack{|n|<2|m|\\n\neq m}} \Big(\frac{|\beta_n|^{8D}}{|\beta_n-\alpha_m|} + \frac{1}{|\beta_n|^{1-8D}}\Big) + O\Big(\frac{1}{|\beta_m|^{1-8D}}\Big)$$

$$= O(|\alpha_m|^{8D-1}).$$

Consequently

$$\sum_{-N}^{N} \frac{|f(\beta_n)|}{|H'(\beta_n)||\beta_n|} \leqq \sum_{-\infty}^{\infty} |A_m||\alpha_m|^{12D-1}$$

$$\leqq \Big(\sum_{-\infty}^{\infty}|A_m|^2 \sum_{-\infty}^{\infty}|\alpha_m|^{24D-2}\Big)^{1/2}$$

$$< \infty$$

since $D<1/24$. Hence $\sum_{-N}^{N} \frac{|f(\beta_n)|}{|H'(\beta_n)|}\frac{1}{|\beta_n|}$ is bounded for all N and so $\sum_{-\infty}^{\infty} \frac{|f(\beta_n)|}{|H'(\beta_n)|}\frac{1}{|\beta_n|}$ is convergent. Hence from above

$$f(z) = H(z)\Big[d_1 + \lim_{N\to\infty}\sum_{-N}^{N} \frac{1}{H'(\beta_n)}\frac{f(\beta_n)}{z-\beta_n}\Big]. \tag{4.29}$$

Now from (2.31) it follows that

$$|f(iy)| \leqq |G(iy)| \sum_{-\infty}^{\infty} \frac{|A_n||\alpha_n|^{4D}}{\sqrt{y^2+\alpha_n^2}}$$

$$\leqq O(|y|^{2D}e^{\pi|y|})\Big(\sum_{-\infty}^{\infty}|A_n|^2 \sum_{-\infty}^{\infty} \frac{|\alpha_n|^{8D}}{y^2+\alpha_n^2}\Big)^{1/2}$$

as $|y|\to\infty$. If $\alpha_N \leqq y < \alpha_{N+1}$, then

$$\sum_{-\infty}^{\infty} \frac{|\alpha_n|^{8D}}{y^2+\alpha_n^2} = \sum_{|n|<|N|} \frac{|\alpha_n|^{8D}}{y^2+\alpha_n^2} + \sum_{|n|\geqq|N|} \frac{|\alpha_n|^{8D}}{y^2+\alpha_n^2}$$

$$< \sum_{|n|<|N|} \frac{|\alpha_n|^{8D}}{|\alpha_n|^2} + \sum_{|n|\geqq|N|} \frac{|\alpha_n|^{8D}}{|\alpha_n|^2}$$

$$< \sum_{|n|<|N|} \frac{|\alpha_n|^{8D}}{|\alpha_n|^2} + \frac{1}{|\alpha_N|^{16D}} \sum_{|n|\geqq|N|} \frac{|\alpha_n|^{8D}}{|\alpha_n|^{2-16D}}$$

$$= O\Big(\frac{1}{|y|^{16D}}\Big).$$

Hence for $|y| \to \infty$,

$$f(iy) = O(|y|^{2D} e^{\pi|y|}) O(\frac{1}{|y|^{8D}})$$

$$= O(\frac{1}{|y|^{\frac{1}{4}}} e^{\pi|y|}).$$

Therefore in (4.29) d_1 should vanish.

We shall now apply Theorem 1 to prove Whittaker's result (Theorem G). If $f(z)$ is defined by (2.22) and (2.21) holds then for $x > 0$

$$|f(x)| \leqq \sum_0^{\infty}{}' \frac{|a_n|}{|x-n|} + \sum_0^{\infty} \frac{|a_{-n}|}{|x+n|} + o(\frac{x}{\log x})$$

$$= x \sum_{n<x} \frac{|a_n|}{n} (\frac{1}{x-n} - \frac{1}{x}) + x \sum_{n>x} \frac{|a_n|}{n} (\frac{1}{n-x} + \frac{1}{x})$$

$$+ \sum_0^{\infty} \frac{|a_{-n}|}{|x+n|} + o(\frac{x}{\log x})$$

$$= O(x).$$

Similarly

$$f(x) = O(|x|).$$

Let now

$$F(z) = \frac{f(z) - f(0)}{z}.$$

Clearly, $F(z)$ satisfies (2.2). Besides, for certain positive constants c_{20} and c_{21}

$$\int_{-\infty}^{\infty} |F(x+iy)| e^{-\pi|y|} dy < c_{20} \sum_0^{\infty} \frac{|a_n|}{n} \int_{-\infty}^{\infty} \frac{n\, dy}{\sqrt{x^2+y^2}\sqrt{(x-n)^2+y^2}}$$

$$+ c_{20} \sum_0^{\infty} \frac{|a_{-n}|}{n} \int_{-\infty}^{\infty} \frac{n\, dy}{\sqrt{x^2+y^2}\sqrt{(x+n)^2+y^2}}$$

$$< c_{21} \sum_0^{\infty} \frac{|a_n|}{n} \int_1^{\infty} \frac{n\, dy}{\sqrt{x^2+y^2}\sqrt{(x-n)^2+y^2}}$$

$$+ c_{21} \sum_0^{\infty} \frac{|a_{-n}|}{n} \int_1^{\infty} \frac{n\, dy}{\sqrt{x^2+y^2}\sqrt{(x+n)^2+y^2}}.$$

If $c_{22} > 0$ is sufficiently small and $c_{23} > 0$ is sufficiently large then for $x > 0$

$$\theta(x) = \sum_0^\infty \frac{|a_n|}{n} \int_1^\infty \frac{n\,dy}{y\sqrt{(x-n)^2+y^2}}$$

$$= \Big(\sum_0^{[c_{22}x]} + \sum_{[c_{22}x]+1}^{[x]} + \sum_{[x]+1}^{[c_{23}x]} + \sum_{[c_{23}x]+1}^{\infty}\Big) \frac{|a_n|}{n} \int_1^\infty \frac{n\,dy}{y\sqrt{(x-n)^2+y^2}}$$

$$= O\Big[\sum_0^{[c_{22}x]} \frac{|a_n|}{n} \int_1^\infty \frac{n\,dy}{y\sqrt{y^2+\{(1-c_{22})x\}^2}}$$

$$+ \sum_{[c_{22}x]+1}^{[x]} \frac{|a_n|}{n} \int_1^\infty \frac{[x]}{y^2}\,dy$$

$$+ \sum_{[x]+1}^{[c_{23}x]} \frac{|a_n|}{n} \int_1^\infty \frac{[c_{23}x]}{y^2}\,dy$$

$$+ \sum_{[c_{23}x]+1}^{\infty} \frac{|a_n|}{n} \int_1^\infty \frac{n\,dy}{y\sqrt{\{(c_{23}-1)n\}^2+y^2}}\Big]$$

$$= O\Big[\sum_0^{[c_{22}x]} \frac{|a_n|}{n} \frac{[c_{22}x]}{x} \log x + \sum_{[c_{22}x]+1}^{[x]} \frac{|a_n|}{n}[x]$$

$$+ \sum_{[x]+1}^{[c_{23}x]} \frac{|a_n|}{n}[c_{23}x] + \sum_{[c_{23}x]+1}^{\infty} \frac{|a_n|}{n} \log n\Big]$$

$$= O(x) \sum_0^\infty \frac{|a_n|}{n} \log n$$

$$= O(x).$$

In the same way we can prove that for negative x

$$\theta(x) = O(|x|).$$

It is now easy to see that $F(z)$ satisfies (2.3). In fact, the function $F(z+\lambda)$, $0<\lambda<1$ also satisfies (2.2) and (2.3). Therefore if $\gamma_n = n+\lambda$, then by Theorem 1

$$\frac{f(z)-f(0)}{z} = \tfrac{1}{\pi}\sin\pi(z-\lambda) \lim_{N\to\infty} \sum_{-N}^{N} (-1)^n \frac{f(\gamma_n)-f(0)}{\gamma_n(z-\gamma_n)} e^{-\epsilon|\gamma_n|}.$$

Since $\dfrac{f(\gamma_n)-f(0)}{\gamma_n(z-\gamma_n)} = O\Big(\dfrac{1}{|n|}\Big)$ it follows by Lemma 8 that the series on the right is convergent and therefore

$$f(z)-f(0) = \tfrac{1}{\pi}\sin\pi(z-\lambda)\Big[\sum_{-N}^{N}(-1)^n \frac{f(\gamma_n)-f(0)}{\gamma_n(z-\gamma_n)} z + \delta_N\Big]$$

where $\delta_N \to 0$ as $N \to \infty$. Thus

$$f(z) - \frac{1}{\pi} \sin \pi (z - \lambda) \sum_{-N}^{N} (-1)^n \frac{f(\gamma_n)}{z - \gamma_n}$$
$$= f(0) - \frac{1}{\pi} \sin \pi (z - \lambda) \sum_{-N}^{N} (-1)^n \frac{f(0)}{z - \gamma_n}$$
$$+ \frac{1}{\pi} \sin \pi (z - \lambda) \sum_{-N}^{N} (-1)^n \frac{f(\gamma_n)}{\gamma_n}$$
$$- \frac{1}{\pi} \frac{\sin \pi (z - \lambda)}{\sin \pi \lambda} \sin \pi \lambda \sum_{-N}^{N} (-1)^n \frac{f(0)}{\gamma_n}$$
$$+ \frac{1}{\pi} \delta_N \sin \pi (z - \lambda).$$

As $N \to \infty$, $\frac{1}{\pi} \sin \pi (z - \lambda) \sum_{-N}^{N} (-1)^n \frac{f(0)}{z - \gamma_n}$ and $\frac{1}{\pi} \sin \pi\lambda \sum_{-N}^{N} (-1)^n \frac{f(0)}{\gamma_n}$ both tend to $f(0)$.
Consequently

$$- \sin \pi\lambda \{ f(z) - \frac{1}{\pi} \sin \pi (z - \lambda) \lim_{N \to \infty} \sum_{-N}^{N} (-1)^n \frac{f(\gamma_n)}{z - \gamma_n} \}$$
$$= \sin \pi (z - \lambda) \{ f(0) - \frac{1}{\pi} \sin \pi\lambda \lim_{N \to \infty} \sum_{-N}^{N} (-1)^n \frac{f(\gamma_n)}{\gamma_n} \}. \tag{4.30}$$

But

$$\sum_{-N}^{N} \frac{|f(\gamma_n)|}{|\gamma_n|} \leqq \sum_{-N}^{N} \frac{1}{|\gamma_n|} |\sin \pi\gamma_n| \sum_{-\infty}^{\infty} \frac{|a_\nu|}{|\gamma_n - \nu|}$$
$$\leqq \sum_{-\infty}^{\infty} |a_\nu| \sum_{-N}^{N} \frac{1}{|\gamma_n| |\gamma_n - \nu|},$$

and

$$\sum_{|n| \leqq N} \frac{1}{|\gamma_n| |\gamma_n - \nu|} = \Big(\sum_{|n| < 2|\nu|} + \sum_{2|\nu| \leqq |n| \leqq N} \Big) \frac{1}{|\gamma_n| |\gamma_n - \nu|}$$
$$< \frac{1}{|\nu|} \sum_{|n| < 2|\nu|} \frac{|\gamma_n| + |\gamma_n - \nu|}{|\gamma_n| |\gamma_n - \nu|} + O(1) \sum_{\nu}^{\infty} \frac{1}{|\gamma_n|^2}$$
$$< \frac{1}{|\nu|} \sum_{|n| < 2|\nu|} \Big(\frac{1}{|\gamma_n - \nu|} + \frac{1}{|\gamma_n|} \Big) + O\Big(\frac{1}{\nu}\Big)$$
$$= O(1) \Big\{ \frac{\log(|\nu| + 1)}{|\nu|} \Big\}.$$

Hence $\sum_{-N}^{N} \frac{|f(\gamma_n)|}{|\gamma_n|}$ is bounded for all N and so $\sum_{-\infty}^{\infty} \frac{|f(\gamma_n)|}{|\gamma_n|}$ is convergent. Thus the limit on the right of (4.30) exists and therefore does that on the left. Consequently

$$f(z) = \sin \pi (z - \lambda) \{ d' + \frac{1}{\pi} \sum_{-\infty}^{\infty} (-1)^n \frac{f(\gamma_n)}{z - \gamma_n} \}$$

where d' is a constant. From (2.22) it follows that $f(iy) = O\left(\frac{e^{\pi|y|}}{|y|}\right)$ as $|y| \to \infty$. This is possible only if d' is zero. This proves the desired result.

Proof of Theorem 12. The condition $D < \frac{1}{4}$ ensures that λ_0 is the λ_n closest to zero and therefore (3.3′), (3.3″) and (3.4)

$$\begin{aligned}
|f(x)| &= \Big| G(x) \sum_{-\infty}^{\infty} \frac{A_n}{(x-\alpha_n)G'(\alpha_n)} \Big| \\
&= O(|x|^{4D}) \sum_{-\infty}^{\infty}{}' \frac{|A_n|\,|\alpha_n|^{4D}}{|x-\alpha_n|} + O(|x|^{8D}) \\
&\leqq O(|x|^{4D}) \Big(\sum_{-\infty}^{\infty} |A_n|^2 |\alpha_n|^{4D} \sum_{-\infty}^{\infty}{}' \frac{|\alpha_n|^{4D}}{|x-\alpha_n|^2} \Big)^{1/2} + O(|x|^{8D}) \\
&\leqq O(|x|^{8D}).
\end{aligned}$$

Let now

$$F(z) = \frac{f(z)-f(0)}{z}.$$

Clearly $F(z)$ satisfies (2.2). Besides

$$\begin{aligned}
&\int_{-\infty}^{\infty} (1+|y|)^{2D} |F(x+iy)|\, e^{-\pi|y|}\, dy \\
&= O(1) \int_1^{\infty} (1+y)^{2D} \frac{|f(z)|}{|z|} e^{-\pi y}\, dy \\
&= O(1) \int_1^{\infty} (1+y)^{2D} \frac{|z|^{4D}}{y^{2D}|z|} \sum_{-\infty}^{\infty} \frac{|A_n|\,|\alpha_n|^{4D}}{\sqrt{y^2+(x-\alpha_n)^2}}\, dy \\
&= O(|x|^{4D}) \int_1^{\infty} (1+y)^{2D} \frac{y^{2D}}{y} \sum_{-\infty}^{\infty} \frac{|A_n|\,|\alpha_n|^{4D}}{\sqrt{y^2+(x-\alpha_n)^2}}\, dy \\
&= O(|x|^{4D}) \int_1^{\infty} \sum_{-\infty}^{\infty} \frac{|A_n|\,|\alpha_n|^{4D}}{y^{1-4D}\sqrt{y^2+(x-\alpha_n)^2}}\, dy \\
&= O(|x|^{4D}) \sum_{-\infty}^{\infty}{}' \frac{|A_n|\,|\alpha_n|^{4D}}{|x-\alpha_n|} + O(|x|^{4D}) \\
&= O(|x|^{4D}) \Big(\sum_{-\infty}^{\infty} |A_n|^2 |\alpha_n|^{4D} \sum_{-\infty}^{\infty}{}' \frac{|\alpha_n|^{4D}}{|x-\alpha_n|^2} \Big)^{1/2} + O(|x|^{4D}) \\
&= O(|x|^{8D}).
\end{aligned}$$

Thus $F(z)$ also satisfies (2.3). Therefore from Theorem 1, we have

$$\frac{f(z)-f(0)}{z} = H(z) \lim_{\epsilon\to 0} \sum_{-\infty}^{\infty} \frac{1}{H'(\beta_n)} \frac{f(\beta_n)-f(0)}{\beta_n(z-\beta_n)} e^{-\epsilon|\beta_n|}$$

or

$$f(z)-f(0)=zH(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\beta_n H'(\beta_n)}\frac{f(\beta_n)-f(0)}{z-\beta_n}e^{-\epsilon|\beta_n|}.$$

But

$$f(0)=H(z)\frac{f(0)}{H(0)}+zH(z)\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{\beta_n H'(\beta_n)}\frac{f(0)}{z-\beta_n}e^{-\epsilon|\beta_n|}.$$

Therefore

$$f(z)=H(z)\{\frac{f(0)}{H(0)}+\lim_{\epsilon\to 0}\sum_{-\infty}^{\infty}\frac{1}{(z-\beta_n)H'(\beta_n)}\frac{f(\beta_n)}{\beta_n}e^{-\epsilon|\beta_n|}\}.$$

UNIVERSITY OF MONTREAL,
MONTREAL, CANADA.

REFERENCES.

[1] N. I. Ahiezer, "On the interpolation of entire transcendental functions of finite order," *Doklady Akademii Nauk SSSR* (N. S.), vol. 65 (1949), pp. 781-784 (Russian).

[2] S. N. Bernstein, "The extension of properties of trigonometric polynomials to entire functions of finite degree," *Izvestiya Akademii Nauk SSSR,* Ser. Mat., vol. 12 (1948), pp. 421-444 (Russian).

[3] V. Bernstein, *Leçons sur les progrès récents de la théorie des séries de Dirichlet,* Gauthier-Villars, Paris, 1933.

[4] ———, "Géneralisation et conséquences d'un Théorème de Le Roy-Lindelöf," *Bulletin des Sciences Mathematiques* (2), vol. 52 (1928), pp. 420-436.

[5] R. P. Boas, Jr., *Entire Functions,* Academic Press, New York, 1954.

[6] ———, "Entire functions bounded on a line," *Duke Mathematical Journal.* vol 6 (1940), pp. 148-169.

[7] ———, "Inequalities between series and integrals involving entire functions," *Journal of the Indian Mathematical Society* (N. S.), vol. 16 (1952), pp. 127-135.

[8] ———, "Growth of analytic functions along a line," *Journal d'Analyse Mathematique,* vol. 4 (1954/1955), pp. 1-28.

[9] R. P. Boas, Jr. and A. C. Schaeffer, "A theorem of Cartwright," *Duke Mathematical Journal,* vol. 9 (1942), pp. 879-883.

[10] M. L. Cartwright, "On certain integral functions of order 1," *Quarterly Journal of Mathematics, Oxford Series* (1), vol. 7 (1936), pp. 45-55.

[11] R. J. Duffin and A. C. Schaeffer, "Power series with bounded coefficients," *American Journal of Mathematics,* vol. 67 (1945), pp. 141-154.

[12] W. L. Ferrar, "On the consistency of the cardinal function of interpolation," *Proceedings of the Royal Society of Edinburgh,* vol. 47 (1927), pp. 230-242.

[13] W. H. J. Fuchs, "A generalization of Carlson's theorem," *Journal of the London Mathematical Society,* vol. 21 (1946), pp. 106-110.

[14] G. H. Hardy, "On two theorems of F. Carlson and S. Wigert," *Acta Mathematica*, vol. 42 (1920), pp. 327-339.

[15] J. Korevaar, *Approximation and intepolation applied to entire functions*, Thesis, University of Leiden, 1949.

[16] B. Ya Levin, "On funcions of finite degree, bounded on a sequence of points," *Doklady Academii Nauk SSSR* (N. S.), vol. 65 (1949), pp. 265-268.

[17] N. Levinson, *Gap and density theorems*, American Mathematical Society, New York, 1940.

[18] ———, "On certain theorems of Pólya and Bernstein," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 702-706.

[19] A. J. Macintyre, "Laplace's transformation and integral functions," *Proceedings of the London Mathematical Society* (2), vol. 45 (1938), pp. 1-20.

[20] M. E. Noble, "The consistency of cardinal series," *Proceedings of the Cambridge Philosophical Society*, vol. 50 (1954), pp. 139-142.

[21] A. Pfluger, "On analytic functions bounded at the lattice points," *Proceedings of the London Mathematical Society* (2), vol. 42 (1936), pp. 305-315.

[22] G. Pólya, "Untersuchungen über Lucken und Singularitäten von Potenzreihen," *Mathematische Zeitschrift*, vol. 29 (1929), pp. 549-640.

[23] E. C. Titchmarsh, *The theory of functions*, 2d ed. Oxford University Press, 1939.

[24] ———, "The zeros of certain integral functions," *Proceedings of the London Mathematical Society* (2), vol. 25 (1926), pp. 283-302.

[25] G. Valiron, "Sur la formule d'interpolation de Lagrange," *Bulletin des Sciences Mathematiques*, vol. 49 (1926), pp. 181-192, 203-224.

[26] J. M. Whittaker, *Interpolatory function theory* (Cambridge Tracts in Mathematics, 1935).

[27] D. V. Widder, *The Laplace transform*, Princeton University Press, 1946.

AMERICAN

JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXXXVII

1965

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND 21218
U. S. A.

INDEX